(2)

NAVAL POSTGRADUATE SCHOOL
Monterey, California

AD-A257 508

DTIC
S ELECTE DEC 01 1992 D
A

THESIS

10
92-30436

MEASURING NAVSPASUR SENSOR PERFORMANCE
USING LOGISTIC REGRESSION MODELS

by

Brook R. Roberts

September, 1992

Thesis Advisor: Professor So Young Sohn

Approved for public release; distribution is unlimited.

92 11 ... 073

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b. OFFICE SYMBOL *(If applicable)* 55	7a. NAME OF MONITORING ORGANIZATION Naval Postgraduate School
6c. ADDRESS *(City, State, and ZIP Code)* Monterey, CA 93943-5000	7b. ADDRESS *(City, State, and ZIP Code)* Monterey, CA 93943-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL *(If applicable)*	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS *(City, State, and ZIP Code)*	10. SOURCE OF FUNDING NUMBERS Program Element No. / Project No. / Task No. / Work Unit Accession Number

11. TITLE *(Include Security Classification)*
MEASURING NAVSPASUR SENSOR PERFORMANCE USING LOGISTIC REGRESSION MODELS

12. PERSONAL AUTHOR(S) Brook R. Roberts

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT *(year, month, day)*	15. PAGE COUNT
Master's Thesis	From To	September 1992	70

16. SUPPLEMENTARY NOTATION
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17. COSATI CODES			18. SUBJECT TERMS *(continue on reverse if necessary and identify by block number)*
FIELD	GROUP	SUBGROUP	

19. ABSTRACT *(continue on reverse if necessary and identify by block number)*

Since its establishment the Naval Space Surveillance Command (NAVSPASUR) in Daihgren Virginia has been providing surveillance data (NAVSPASUR data sets) for thousands of space objects in a near earth orbit. To date, very little statistical analysis of these data sets in the form of a system performance evaluation has been conducted.

The objective of this thesis is to provide NAVSPASUR with a statistical method to evaluate the system performance in terms of its capability of detecting space objects.. In this thesis six individual station models, as well as a system-wide model are estimated. Optimal probability levels for classifying predictions are additionally provided. The results being provided are obtained through the implementation of Logistic Regression analysis. The system-wide model estimated in this thesis, is superior in its prediction accuracy when compared to the previous model provided to NAVSPASUR in a September 1991, Naval Postgraduate School Master's Thesis. Finally an implementation program written in FORTRAN is given. This program provides a user friendly interface capability for predicting system performance in terms of its detection ability.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS REPORT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL Brook R. Roberts	22b. TELEPHONE *(Include Area code)* (408) 646-2772	22c. OFFICE SYMBOL 59

Approved for public release; distribution is unlimited.

Measuring NAVSPASUR Sensor Performance Using Logistic Regression Models

by

Brook R Roberts
Lieutenant, United States Navy
B.S., Maine Maritime Academy 1984

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN SPACE SYSTEMS TECHNOLOGIES
(SPACE SYSTEMS OPERATIONS)
from the
NAVAL POSTGRADUATE SCHOOL
September 1992

Author: Brook R Roberts
Brook R. Roberts

Approved by:
Professor So Young Sohn, Thesis Advisor

Dan C Boger
Professor Dan Boger, Second Reader

R Panholzer
Professor Roudolf Panholzer, Chairman Space Systems
Academic Group

ABSTRACT

Since its establishment the Naval Space Surveillance Command (NAVSPASUR) in Dalhgren Virginia has been providing surveillance data (NAVSPASUR data sets) for thousands of space objects in a near earth orbit. To date, very little statistical analysis of these data sets in the form of a system performance evaluation has been conducted.

The objective of this thesis is to provide NAVSPASUR with a statistical method to evaluate the system performance in terms of its capability of detecting space objects. In this thesis six individual station models, as well as a system-wide model are estimated. Optimal probability levels for classifying predictions are additionally provided. The results being provided are obtained through the implementation of Logistic Regression analysis. The system-wide model estimated in this thesis, is superior in its prediction accuracy when compared to the previous model provided to NAVSPASUR in a September 1991, Naval Postgraduate School Master's Thesis. Finally an implementation program written in the FORTRAN is given. This program provides a user friendly interface capability for predicting system performance in terms of it detection ability.

DTIC QUALITY INSPECTED 3

Accesion For	
NTIS CRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

TABLE OF CONTENTS

I. INTRODUCTION 1

A. THESIS OBJECTIVES 1

B. THESIS ORGANIZATION 3

II. BACKGROUND 5

A. RADAR FENCE PERFORMANCE RELATED VARIABLES 5

1. Variables Obtained by Radar Fence Collection Elements 5

2. Geomagnetic and Solar Data and their effects. 8

B. RADAR FENCE DESCRIPTION AND THEORY 14

1. Basic system description. 14

2. Basic Radar Theory 16

III. LOGISTIC REGRESSION 19

A. MODEL 19

B. VARIABLE SELECTION 20

C. ESTIMATION 21

D. CROSS VALIDATION 22

IV. DATA ANALYSIS 24
A. PRELIMINARY ANALYSIS 24
1. Data Set Structure 24
2. Descriptive Statistics 26
B. INDIVIDUAL STATION MODELS 26
1. Model Formulation 26
2. Model Cross Validation 29
C. SYSTEM-WIDE MODEL 35
1. Model Formulation 36
2. Cross Validation and Comparisons 37

V. CONCLUSIONS 41

APPENDIX A 43
A. FORTRAN CODING AND RELATED OUTPUT 43
1. Data Set Manipulation Programs 43
2. Fortran Implementation Program 46
B. SAS PROGRAMMING AND RELATED OUTPUT 54
C. FINAL DATA SET USED IN ANALYSIS 61

LIST OF REFERENCES 62

INITIAL DISTRIBUTION LIST 63

I. INTRODUCTION

A. THESIS OBJECTIVES

The Naval Space Surveillance Command (NAVSPASUR) located in Dahlgren, Virginia, is the current operating custodian for a radar fence consisting of three transmitters and six receivers. This fence, operating for over thirty years, has provided the Department of Defense with a unique satellite surveillance capability. What makes this system unique is its ability to acquire and catalog orbital characteristics (stored in a NAVSPASUR data set) for a multitude of earth orbiting objects with virtually no requirement for pre-targeting or cooperation from these objects. Up until the fall of 1991 virtually no statistical analysis techniques, based upon cataloged characteristics, were used to provide a measure of radar fence system performance. Such a model, if accurate, would provide a measurement of effectiveness for the system's performance. The measure of system performance should be based upon the system's ability to detect an object with a given set of orbital characteristics.

In the fall of 1991, LT Schaaf of the Naval Postgraduate School provided NAVSPASUR with a statistical model [Ref. 1: p.31]. The analysis performed was based upon logistic regression. The model provided was expected to predict the probability of detection for a satellite with known orbital characteristics. The parameter estimates of the logistic regression model were based upon a one day data set provided by NAVSPASUR. Results of a cross validation of this model indicated that

there were many non-detections improperly classified as detection. This leads one to question its prediction accuracy. In addition, this analysis did not clearly state the role of the predicted probability of detection in determining the future detection or non-detection capability of a satellite of interest. In other words, no threshold value for classification was provided. Furthermore, that model focused solely upon the overall system performance; it did not analyze the performance of the individual receivers, which is of interest to NAVSPASUR.

The main goal of this thesis is to provide an improved prediction model for system performance. It is also intended to provide individual prediction models for the six receiving stations. The analysis of new logistic regression models is based upon eight days worth of data provided by NAVSPASUR. Additionally, solar, geomagnetic and orbital data, which was not previously analyzed, is incorporated. Further analysis to determine the probability level for each model is also performed. This threshold value is then used to classify the predicted probability of detection as either an actual detection or non-detection. Once all the analysis is completed and the appropriate models are selected, implementation procedures are provided in a FORTRAN program. The program allows the user to determine the probability of detection for a satellite with known orbital characteristics for each of the six receiving stations and for the entire system.

B. THESIS ORGANIZATION

Chapter II provides the reader with a description of how those variables used in the analysis are physically related to the radar fence performance. This chapter also furnishes a description of the radar fence's construction or physical layout. The last section of Chapter II provides a simplified example of the applicable theory of radar operations.

Chapter III gives a description of logistic regression, along with the necessary variable selection, estimation and cross validation procedures. The first two procedures (backward elimination and estimation) are used to select those orbital characteristics which are influential to system performance and are used to estimate the corresponding parameters. The latter can be used to generate the classification table. This table is used as a cross validation tool for the seven final fitted models. The classification table is additionally utilized to determine the threshold value at which predictions are classified as either detections or non-detections.

In Chapter IV, data analysis is performed. Based upon the methodology described in Chapter III, seven final models (six for the individual receivers and one system- wide) are selected, and the results are discussed. The system-wide model selected in this thesis is compared to the previous model, and all tradeoffs are discussed.

Chapter V contains conclusions and recommendations for further study.

Appendix A includes all SAS and FORTRAN code used in the analysis, as well as a brief amount of their related output. A brief description of each program is provided. The implementation program is written in FORTRAN code. This program will be used to

compute the probability of detection of a satellite when the associated information regarding the orbital and solar/geomagnetic characteristics is given. In addition a predetermined threshold value is provided. This value is used to determine whether a satellite with known orbital, geomagnetic, and solar characteristics can be detected or not.

II. BACKGROUND

Chapter II is divided into two sections. Section A provides an explanation of the physical relationship between radar fence performance and the orbital, geomagnetic, and solar characteristics contained within the NAVSPASUR data sets. Section B provides a description of the radar fence physical design. Additionally, it furnishes background discussion of the radar theory applicable to system performance.

A. RADAR FENCE PERFORMANCE RELATED VARIABLES

Variables related to the NAVSPASUR radar fence performance are obtained from two distinct data sets provided by NAVSPASUR. They are the radar fence collection elements data set and the geomagnetic and solar data set.

1. Variables Obtained by Radar Fence Collection Elements

The NAVSPASUR fence and its data collection elements constitute a system that lends itself to statistical analysis. Orbital characteristics for each satellite are observed, on an average, four to seven times daily. These observations provide a significant base of data from which overall system performance can be predicted. Each one of these orbital characteristics is in some way directly or indirectly related to the overall radar fence system performance. Among them the following seven orbital characteristics are discussed:

1. Radar Cross Section (RCS): the cross sectional area in square meters of an object from which radar energy is reflected. The larger the radar cross section, the greater the reflection area and the higher the probability of detection.

2. Orbital Eccentricity: the measure of an orbit's departure from that of a circular orbit. All objects orbiting the earth follow distinct paths. These paths, which vary from case to case, are all members of a family of conic sections referred to as ellipsoids. By definition, an ellipse is a plane curve such that the sum of the distances of each point in its periphery from two fixed points, the foci, are equal [Ref. 2]. The eccentricity of the ellipse is the measure of the distance between the center of the ellipse to either focus. For instance a circle is an ellipse with eccentricity equal to zero. The eccentricity of an orbit will remain constant throughout the orbit. Two satellites with distinctly different orbital eccentricities will follow different orbital paths when traveling through the radar fence's energy field. For example, a satellite with one eccentricity may be ascending when passing through the field, while another with a different eccentricity may be travelling through the field horizontally. Additionally one satellite may pass through the center of the field while another may pass through its edges. A good example of this can be seen by comparing the two orbital paths shown in Figure 2.1. These variations in satellite paths have a direct effect on the detection capability of the radar fence. It should be noted that the radar

fence energy is not uniform throughout. The fence is much weaker at its edges than it is in its center. This alone makes the orbital path critical to detection capability.

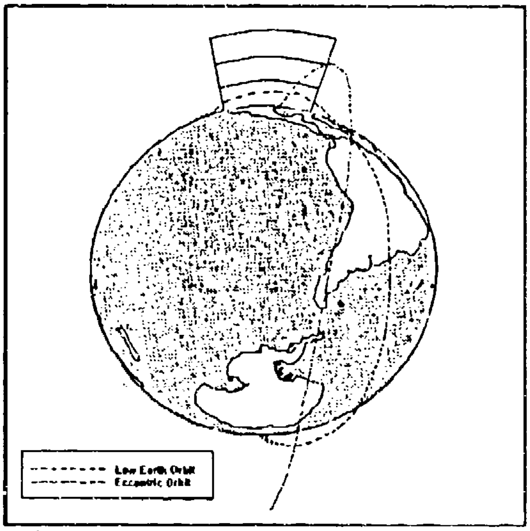

Figure 2.1 Orbital Comparisons

3. Orbital Inclination: the angular measure (in degrees) between the angular momentum vector of the satellite and an axis passing through the center of the earth extending through the north pole. A satellite's orbital inclination determines the path which it will travel over the surface of the earth (i.e., equatorial or polar orbiting). Variations in inclination from satellite to satellite can account for the variations in the orbital paths followed by these satellites when passing through the fence's energy field. Since the energy field of the fence is not uniform throughout, there will once again be fluctuations in the detection capability.

4. Predicted Altitude: the predicted distance in nautical miles from the surface of the earth to the object. The altitude prediction is made for that point in the satellite's orbital path where radar fence energy concentration is the greatest. The higher the object is above the earth the harder it is to detect. Radar power drops off at a rate of one over altitude to the fourth power.

5. Longitude: the longitude at which the satellite entered the radar fence energy field. As stated before the energy field is not constant throughout, so detections tend to fall near the coastal regions of the United States.

6. Orbital period: the time that it takes a satellite to make one complete revolution around the earth. The greater a satellite's altitude, the greater the orbital period. Orbital period is important because it determines the amount of time an object spends in the radar fence's energy field. The more time spent in the field the greater the probability of detection.

7. Latitude: the latitude in degrees north where the satellite is detected by the radar fence. This is important for the same reason as longitude.

2. Geomagnetic and Solar Data and their effects.

The NAVSPASUR command, in addition to collecting the data mentioned above, also receives and maintains a database of solar and geomagnetic data from the United States Air Force. Both solar and geomagnetic anomalies potentially could have

a negative effect on radar fence performance. A brief description of this effect will be discussed later in this chapter. The following variables are considered:

1. Solar Flux: the measurement of the intensity of electromagnetic radiation (including radio waves) emitted by the sun. The intensity of any electromagnetic radiation (including radio waves) is measured as a flux, i.e., in term of energy per unit area per unit time. In general an electromagnetic disturbance in the sun represents a spectrum of waves of all frequencies. Therefore, to determine the total flux over all frequencies one must integrate with respect to frequency using the following equation [Ref. 3: p. 3]:

$$\phi = \int_0^\infty \phi_\upsilon \, d\upsilon$$

where

υ = frequency

ϕ_υ = flux at frequency υ.

The units of Φ_υ are those of energy per area per second per hertz. For the purpose of reporting Φ_υ is taken as 10^{-22} watts per square meter per hertz. Thus a 10.7 cm solar flux index of 140 represents a solar radio flux with wavelength 10.7 cm of 1.4×10^{-20} watts per square meter per hertz. The choice of this parameter (10.7 cm) is predicated upon availability (it is one of the frequencies continuously monitored by various sensors) and the fact that it has been judged to be well correlated with variations in the upper atmosphere [Ref. 3: p. 3].

2. Mean Solar Flux: the mean solar flux for the last 90 days proceeding that date last measured.

3. Mean daily geomagnetic index: the value for the geomagnetic index for the closest day preceding epoch. The official index of the earth's geomagnetic activity is called the Goettigen index [Ref. 3: p. 3]. The index is based upon the measurement of twelve stations around the globe. This information unfortunately is not available in real time. For this reason Air Force Grand Weather Central (from which NAVSPASUR obtains its value of the index) attempts to compute a similar value in real time using six stations of its own. The index is simply an indicator of the general level of activity in the geomagnetic field of the earth. Variations are mostly caused by fluctuations in the strength of the solar wind. Every three hours each of these stations records the difference between the highest and lowest magnetic field strengths measured in that period and reports it as the "range" or "amplitude" for that period [Ref. 3: p. 3]. The reporting observatory assigns a digit between 0 and 9 for each three hour interval to each one of the three field components (x,y,z or north-south, east-west, and vertical up-down, respectively). The amplitude recorded at each station represents the local activity and is found to depend strongly on the geomagnetic latitude of the observatory. It is desirable to remove any latitude-dependence from the data in order to be able to make a direct comparison of the data from the different stations. Therefore, each station applies a correction factor to its data. By

doing this, on average, the stations will tend to report similar values of the amplitude at the same time, however there will still be differences due to the local irregularities. The index of overall global activity, called the geomagnetic planetary index, is the result of averaging the values obtained from each of the six stations [Ref. 3: p. 3]. The mean daily value is computed by averaging the eight three hour intervals recorded.

4. Three hour average geomagnetic indices: simply the values recorded for the eight three hour blocks described in the mean daily geomagnetic index section.

Solar radiation in the form of the solar wind is emitted in all directions from the sun into space. The solar wind, a neutral plasma of negatively and positively charged ions, is emitted through the thermal nuclear expansion of the sun's coronal layer. The solar wind travels through space at a velocity exceeding mach eight [Ref. 4: p. 45]. Eventually it comes into contact with a region of the earth's magnetic field referred to as the magnetosphere. This high speed collision between the solar particle flux and the magnetic field causes a shock wave to form at a altitude of about 15 earth radii above the surface [Ref. 4: p. 45] . This shock is referred to as a bow shock [See Figure 2.2][Ref.5: p. 59]. The actual altitude of formation varies with the force of the solar wind. Behind the bow shock a laminar flow section forms and is referred to as the magnetopause. Another region called the magnetosheath forms between the magnetopause and the bow shock. This is a region of rather distorted

magnetic field, intermixed with irregularly distributed plasma. Within the magnetosphere, the magnetic field dominates the motion of the charged particles of the plasma. Belts of charged particles called the Van Allen Belts are a product of this domination. Below the Van Allen belts is the region referred to as the plasmasphere. The plasmasphere forms the lower boundary of the magnetosphere with a region called the ionosphere [Ref. 5: p. 59].

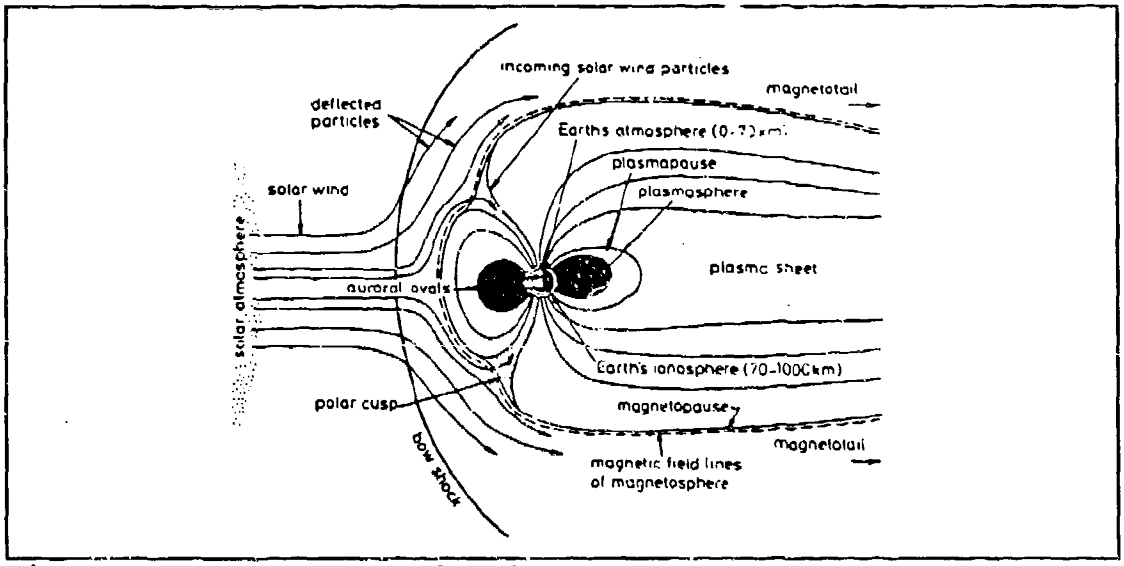

Figure 2.2 Earth Magnetic Field and Atmosphere

The ionosphere is a region of ionized plasma that extends from approximately 50 km to 2000 km above the earth's surface. The ionization of the atmospheric particles within this region is caused by the electromagnetic radiation contained in the solar wind. Only a portion of the molecules within this region are actually ionized. The magnetic field of the earth interacts with the ionized particles of the ionosphere and aligns them with the field strength pattern of the earth [Ref. 5]. Sunlit portions of the earth's

atmosphere receive a stronger flux of sun-born particles, and are consequently more ionized than areas that are not in direct contact. This characteristic causes the ionosphere to expand during daylight hours and collapse during the night. In studying the ionosphere, scientists have determined that fluctuations, caused by changing solar and geomagnetism effects, have direct influence on how radio waves travel through the earth's atmosphere. Free electrons, created by the ionization in earth's ionosphere, can greatly affect a radio wave's propagation (ability to travel through atmosphere). These free electrons are capable of absorbing incident radio wave energy at any radio wave frequency. Given this, one can deduce that the amount of absorption, reflection, or refraction that a radio wave experiences is related to both the radio wave's frequency and the concentration of electrons in the atmosphere. Theoretically, if electron concentration in the ionosphere reached a high enough level, radio waves could be greatly refracted or even reflected back to earth. From this, the following formula was deduced.

$$f_c = 8.9788 * 10^{-6}\sqrt{N}$$

where
f_c = *critical frequency*
N = *number of electrons per cubic meter.*

Any radio wave at a frequency below the critical frequency is going to be refracted to a such a degree that it effectively will be reflected back by the ionosphere. The critical frequency fluctuates with seasons, ionized region, and time of day.

Both the geomagnetic index and solar flux are directly related to the changes in the atmosphere discussed above. If the solar flux increases, so will the total electron

concentration in the ionosphere. If there are fluctuations in the solar wind, there will be fluctuations in the geomagnetic field. This in turn causes fluctuations in the earth's regional geomagnetic index. For these reasons, the geomagnetic index and solar flux are considered to have a direct physical connection to the radar fence performance.

B. RADAR FENCE DESCRIPTION AND THEORY

1. Basic system description.

As mentioned previously the radar fence system consists of three transmitting and six receiving stations or sites. The three transmitting sites, located in Lake Kickapoo TX, Gila River AZ and Jordan Lake AL, are positioned on a great circle, stretching across the southern United States [See Figure 2.3]. Each of the three transmitting sites operates a transmitting antenna consisting of a linear array of dipole elements aligned north to south. The antennas transmit a continuous-wave signal at a frequency of 216.980 MHz. The largest of the transmitters, located in Lake Kickapoo TX, operates at 810 kW of output power and consists of eighteen separate collinear bays stretching 3200 meters from north to south. The transmitter site is separated into two distinctly separate components referred to as the North and South transmitters. They may both be operated individually if required.

The remaining two transmitters, Gila River and Jordan Lake, operate at a power output of 45 kW and consist of single bay antenna arrays. The Gila River site is 484 meters long with 384 elements. The Jordan river site is 311 meters long with 256

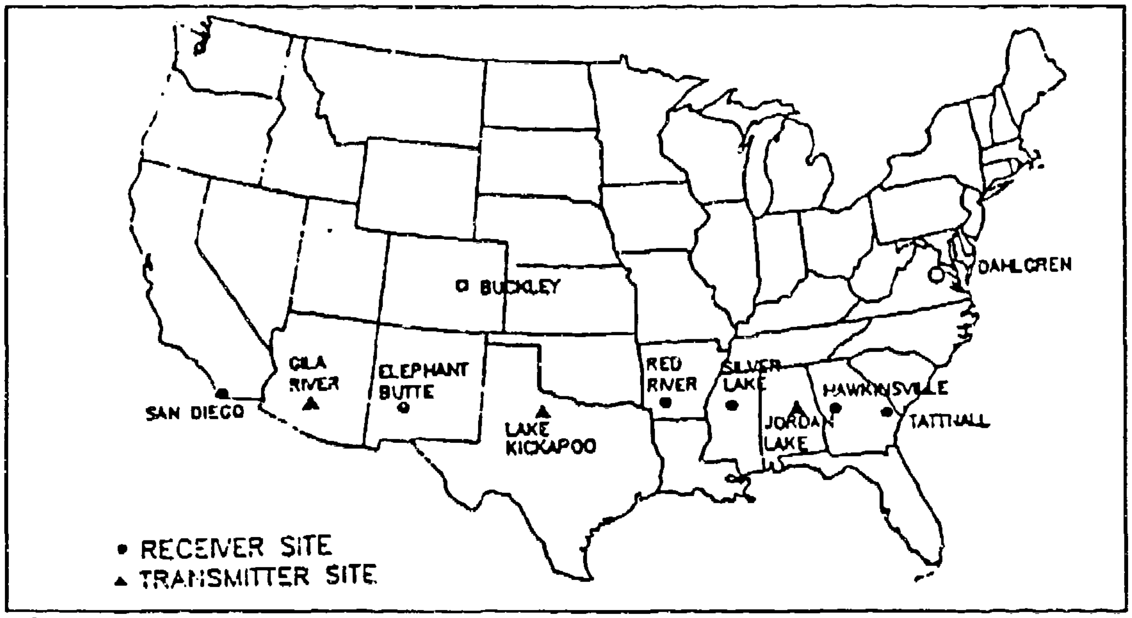

Figure 2.3 Radar Fence Layout

elements.

The six receiving sites are located in San Diego CA, Elephant Butte NM, Red River AR, Silver Lake MS, Hawkinsville GA, and Tattnal GA. The San Diego receiving station operates an antenna layout with a plus (+) configuration, consisting of eight linear dipole arrays extending east to west and four extending north to south. Each array is 600 feet long and is perpendicular to the fence plane which is inclined 33 degrees with respect to the equator. The San Diego receiver is slated for conversion to the new Saint Andrews cross configuration. This configuration is simply a pattern formed by connecting the diagonals of a square [Ref. 6]. The Elephant Butte receiving station is currently being brought back on line as a Saint Andrews cross configuration. The Saint Andrews cross configuration allows for high altitude tracking. High altitude station arrays are 2400 feet long. There are ten arrays which are deployed along lines

rotated 45 degrees with respect to the fence plane. The Hawkinsville receiving station has the same configuration as Elephant Butte. The remaining two stations are low altitude stations similar to the San Diego receiving station previously mentioned. They both have twelve antenna arrays and are also laid out in a Saint Andrews cross configuration.

2. Basic Radar Theory

The three transmitting and six receiving stations working in conjunction with one another form a fan of electromagnetic energy which spans the continental United States. [See Figure 2.2] To simplify the description of the radar theory applicable to the operation of the radar fence, a single unit consisting of one transmitter and one receiver will be used. An independent radar system must have at least one transmitter and one receiver.

Radar theory itself is based upon stochastic, non-deterministic processes. To explain such processes in detail is not only beyond the scope of this thesis, but also is not the intent. It is, on the other hand, important to show how the orbital and atmospheric characteristics discussed earlier in this chapter are related to the physical operation of a radar. Orbital characteristics (i.e., inclination, eccentricity, etc.) all have a direct effect on the range between a satellite and the receiving station. Range has a direct effect on a radar system's capability to detect or not detect an object with a given radar cross section. Atmospheric effects, including geomagnetic and solar anomalies, in the form of losses are very important in the design stages of all radar systems. To best demonstrate how each of these effects are related to radar system performance, a

simplified form of the radar equation is provided. A standard simplified form of the radar equation is [Ref. 7: p. 3]:

$$R^4 = P_t G_t G_r \frac{\lambda^2 \sigma}{4\pi^3 P_r L}$$

where
P_r = *Receiver power (kW)*
P_t = *Transmitter power (kW)*
G_t = *Transmitter antenna gain*
G_r = *Receiver antenna gain*
λ = *Wavelength (m)*
σ = *Radar cross section*
R = *Range (m)*
L = *Losses*

In the above equation all variables, for the purpose of explanation, are held constant except for radar cross section of the object and range to the object (up to the system's maximum theoretical range). The operating parameters for the Lake Kickapoo transmitter and the San Diego receiver will be used. The following parameters are given [Ref. 1: p. 6-7]:

P_r = 6.76 X 10^{-19} kW
P_t = 810 kW
G_t = 10000
G_r = 316.23
λ = 1.38 meters
σ = variable (meters squared)
R_σ = variable (meters)

By plugging these values into the radar equation given above and converting to nautical miles we obtain:

$R_\sigma = 3739.92*\sigma$

A plot of this equation for variation in altitude (in nautical miles) versus radar cross section (meters) is provided [See Figure 2.4]. By analyzing this plot one can determine whether a satellite with a given radar cross section and range is detectable. Any value above the curve will not be detected.

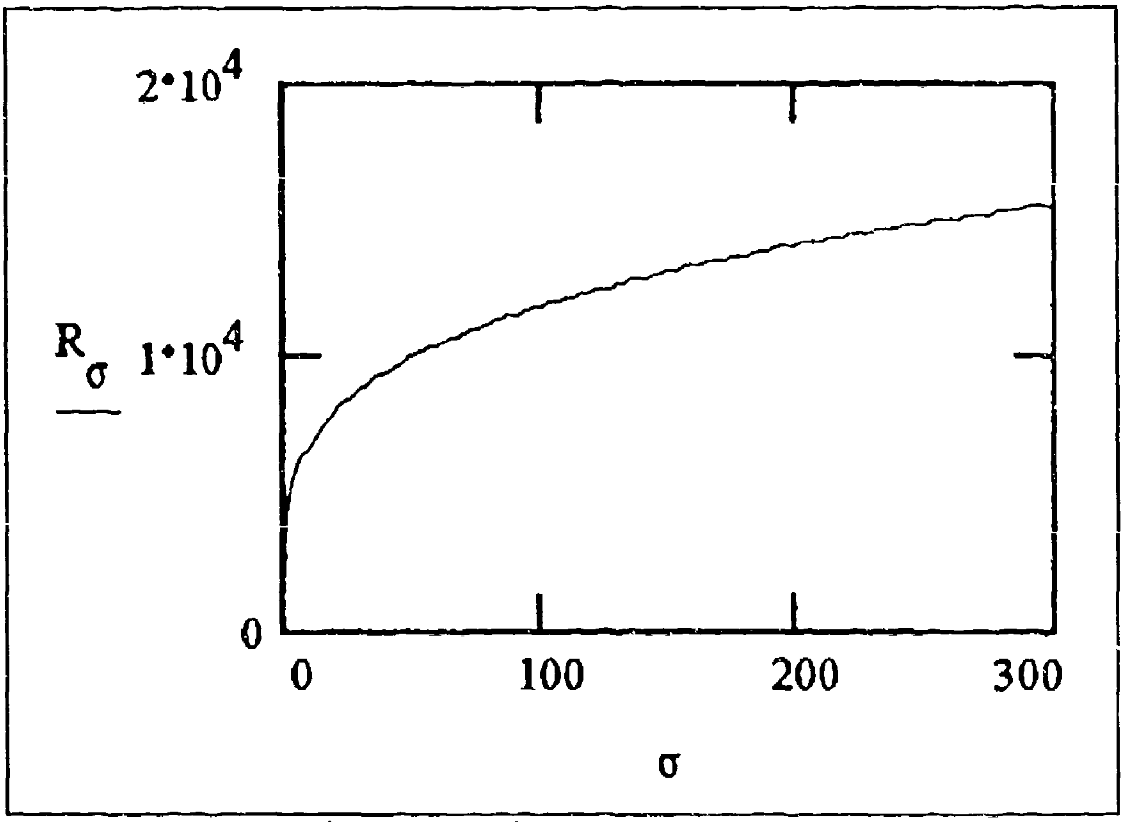

Figure 2.4 Satellite Detection for RCS and Range

III. LOGISTIC REGRESSION

Based on the physical characteristics of a satellite's orbit and the radar theory discussed in Chapter II, it is of interest to NAVSPASUR to formulate prediction models for the performance of the radar fence in terms of its ability to detect a satellite. A response variable, which has two outcomes such as detection or non-detection of a satellite, can be assumed to follow a binomial distribution where the probability of detection can be predicted as a function of associated orbital characteristics of the satellite and other characteristics. A statistical analysis, logistic regression, is widely used to predict such a probability.

In this chapter, logistic regression is introduced, along with the necessary estimation method, variable selection procedure, and cross validation method.

A. MODEL

Logistic regression is often used to relate a probability of occurrence of a categorical (detection/non-detection) outcome to a set of explanatory variables. Once the relationship is established based on the available data, estimated models can be used to predict the future outcome of the categorical response variable, when the values of explanatory variable(s) are given.

Let y_i be the observed number of successes out of n_i independent trials for the ith experiment (i=1,..,n). In logistic regression it is assumed that y_i is a binomial random variable for the ith experiment with n_i trails and a probability of success of θ_i. The value

of θ_i is unknown and is greater than or equal to zero or less than or equal to one. As stated before, in logistic regression θ_i is modeled as a function of predictors or explanatory variables, x_i's using the cumulative distribution of the logistic function [Ref. 8: p. 269]:

$$\theta_i = \frac{\exp(\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\ldots+\beta_q x_{iq})}{1+\exp(\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\ldots+\beta_q x_{iq})}$$

Another equivalent form is:

$$\begin{aligned} logit(\theta_i) &= \ln\left[\frac{\theta_i}{(1-\theta_i)}\right] \\ &= \beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\ldots+\beta_q x_{iq} \end{aligned}$$

B. VARIABLE SELECTION

When it is not clear which subset of the explanatory variables $(x_1,..,x_q)$ has the greatest combined influence on the variation of the response variable, it is necessary to use a variable selection scheme. A generally used method for this is stepwise regression. There are three basic forms of the algorithm used in stepwise regression: forward selection, backward elimination, and stepwise elimination.

For instance, when using the backward elimination procedure all predictors are included in the model and the parameter estimates are determined. Next, a chi-square goodness of fit test for the following hypothesis is performed:

H_0: $\beta_k=0$

H_a: $\beta_k \neq 0$

$(k=1,..,q)$

The variable whose estimate is determined to be the least significant is removed if its p-value is greater that the predetermined significance level. Once the variable is removed the process is repeated with the remaining variables until no further variables meet the requirement for removal. The resulting model is selected for the purpose of the prediction of radar fence performance.

C. ESTIMATION

In this section, estimation procedures for the logistic regression model are briefly explained.

Given that $var(y_i/n_i)=\theta_i(1-\theta_i)/n_i$, it follows that the variances of the binomial response variables may often differ. Hence it would seem appropriate to use weighted least squares estimation of the parameters of the $logit(y_i/n_i)$ model with weights $w_i=1/n_i[\theta_i(1-\theta_i])$[Ref.8: p. 269]. One problem that arises when doing so is that θ_i and hence w_i are unknown. Through the use of an iterative procedure or algorithm (i.e., the Iteratively Re-weighted Least Squares (IRLS) algorithm used in SAS) one can estimate the θ_i's and thus w_i's for the given θ_i's. The process starts first by estimating β_k's that can be used in the logistic regression equation. These β_k's provide an initial estimate of θ_i which will be denoted as θ_{i0}. The initial estimate, θ_{i0}, is then used in the following equation to obtain the adjusted response z_i [Ref:8: p. 269]:

$$z_i = logit(\theta_{io}) + \frac{(y_i - n_i\theta_{io})}{n_i\theta_{io}(1-\theta_{io})}$$

The adjusted response z_i is used in the iteration to compute the maximum likelihood estimates. This is done by setting $w_i = 1/n_i[\theta_{i0}(1-\theta_{i0})]$ and then computing the linear regression of z_i on the predictors using the weights w_i. The resulting estimates of the β_k's are then used again to update the estimates of the θ_i's. These estimates of θ_i's are then used to start the process again. The process will continue until a predetermined stopping criterion is met, resulting in the final estimates of the β's.

The SAS PROC LOGISTIC procedure uses a stopping criterion referred to as a convergence criterion. The iterations are considered to have converged when the maximum change (either relative or absolute) in parameter estimates between successive steps is less than the value specified [Ref. 9: p. 1080]. The default specification in SAS is 1E-4 or .0001. When the maximum change between estimates from successive steps reaches a value less than or equal to 0.0001 the stopping criterion is met, and the remaining estimates are used. A relative change criterion (the ratio of the change in estimate values to the estimate from the previous step) is used if the parameter is greater than 0.01 in absolute value. Otherwise an absolute change is used.

D. CROSS VALIDATION

Once all the parameter estimates are computed and a model is established, the estimated probability of event responses can be obtained by using the variables provided by the original data set. By doing so one introduces an error-count estimate which is

biased. One way of reducing such a error-count bias is through the use of the jackknife procedure. The jackknife procedure provides the analyst with both a cross validation capability as well as a means of classifying predicted responses as actual events or non-events.

The jackknife procedure is used not only to decrease error count bias but also to provide a means of model cross validation. The jackknife procedure accomplishes this by removing the trial to be classified, re-estimating the parameter estimates, and then classifying the trial based on these new parameter estimates. This would be a very costly and time consuming process if this process were to be repeated every time a trial is removed. The LOGISTIC procedure included in the SAS program provides a one-step approximation to obtain the new parameter estimates.

IV. DATA ANALYSIS

The statistical techniques discussed in Chapter III are implemented on the actual data to provide prediction models for the performance of each station as well as that of the entire system. In the first section, descriptive statistics for each of the variables used in the analysis are provided. The second section deals with the results of the estimated detection models for each of the six individual stations. In the third section, the overall system-wide model is predicted, and it is compared to the model suggested by Schaaf [Ref. 1: p.31] in terms of the size of the classification error generated by the two models.

A. PRELIMINARY ANALYSIS

1. Data Set Structure

All the results are based upon the analysis of a randomly selected data set consisting of 47,464 observations from an eight day period (April 20-23 and May 4-7). The data consists of satellite orbital characteristics and geomagnetic/solar measurements which were provided by NAVSPASUR in the form of two distinct data set types (one for solar/geomagnetic and the other for orbital characteristics).

Two FORTRAN programs were written [See Appendix A]. The first program enables one to randomly select trials (between 7,000 and 8,000 per day) from the orbital characteristics data set (containing approximately 38,000 trials per day). Selection was made based upon a variable referred to as satellite catalog number

(SATCAT). Each satellite observed by the radar fence is assigned a satellite catalog number [See Appendix A]. Any of these satellites, which follow some form of an operational orbit, could pass through the energy field of the fence anywhere from five to ten times per day, depending upon their orbital period. Whenever the satellite passes through the fence it is recorded as an observation. Therefore a particular satellite catalog number could be observed multiple times in any given 24 hour period. The FORTRAN program sorts through the data by satellite catalog numbers. The program starts with the first satellite catalog number observed, recording each of the observations until it reaches a new catalog number. Once the new number is reached the program stops and enters a random number generation routine assigning a random number to each observation recorded. Once every observation is assigned a random number the program then determines the observation with the lowest random number and selects it for the output data set. The program starts at the next satellite catalog number and repeats the process. The output data set consists of randomly selected data with at least one observation for each satellite catalog number.

This new data set, based on the randomly selected cases, is then merged with the second data set type (solar and geomagnetic data set) by time and date using the second FORTRAN program [See Appendix A]. These two programs are used for each of the eight days provided. The resulting eight data sets are combined as one final data set using SAS. An example of one observation from the final data set collected on 4 May is provided in Appendix A. The variables used in the analysis include time (HOURS), radar cross section (RCS), orbital eccentricity (ECC), orbital inclination in

radians (INCS), altitude (ALT), altitude squared (ALT2), longitude (LONG), longitude squared (LONG2), orbital period (PER), latitude (LAMBDA), latitude squared (LAMBDA2), daily solar flux (SOLF), longitude*latitude (LLMCROSS), longitude*altitude (LACROSS), altitude*latitude (ALMCROSS), altitude*longitude*latitude (A3CROSS), and three hour geomagnetic index (GEOM). The higher order terms are included to consider possible interaction effects.

2. Descriptive Statistics

The SAS procedure PROC UNIVARIATE [Ref. 10] is used to provide descriptive statistics (mean, standard deviation (STD DEV), first quantile (Q1), median, and third quantile (Q3)) for the variables used in the analysis. The results are provided in Tables 4.1 and 4.2. The full range of variables (except for RCS) are used in the analysis. In view of the interest of NAVSPASUR the values for RCS which exceed 257.2087 were treated as outliers and were not included in the analysis. Additionally, absolute values of RCS are taken in the analysis when they are coded in the data set as negative values. Estimation of the logistic regression models for each of the six individual stations as well as the system-wide model is based upon these values.

B. INDIVIDUAL STATION MODELS

1. Model Formulation

Six separate prediction models for detection are estimated with the SAS PROC LOGISTIC [Ref. 9: p. 1071] for each of the six individual receiving stations. A

TABLE 4.1 DESCRIPTIVE STATISTICS FOR EXPLANATORY VARIABLES

VARIABLE NAME	MEAN	STD DEV
RCS	2.536546	8.269903
ECC	0.094003	0.203543
INC	0.205615	0.315134
ALT	3065.257	6195.36
LONG	99.00969	31.45509
PER	182.5739	208.537
LAMBDA	29.3005	4.986217
SOLF	161.12	30.11862
GEOM	10.45123	3.886088

TABLE 4.2 DESCRIPTIVE STATISTICS FOR EXPLANATORY VARIABLES

VARIABLE NAME	Q1 25%	MEDIAN	Q3 75%
RCS	0.031	0.152	1.974
ECC	0.003	0.0079	0.045
INC	-0.14335	0.197356	0.409432
ALT	857.695	1080.7	1521.17
LONG	73.62	99.11	124.66
PER	102.7	107.6	116.9
LAMBDA	27.6015	30.776	32.793
SOLF	133	173	195
GEOM	6	10	14

copy of the SAS output for receiving station one (San Diego) is provided in Appendix A.

The response profile for each of the six individual stations is given in Table 4.3.

TABLE 4.3 RESPONSE PROFILE FOR INDIVIDUAL STATIONS

STATION	DETECTIONS	NON-DETECT	% DETECT
SAN DIEGO (ONE)	8427	39037	17.75%
ELEPHANT BUTTE (TWO)	6340	41124	13.36%
RED RIVER (THREE)	8881	38583	18.71%
SILVER LAKE (FOUR)	8932	38532	18.82%
HAWKINSVILLE (FIVE)	10005	37459	21.07%
TATTNALL (SIX)	8554	38910	18.02%

These tables contain theactual number of detections and non-detections (from actual input data set) made at each station. In the raw data set provided by NAVSPASUR [See Appendix A] several categories are used to represent different states of detection for each of the individual stations (Y1 through Y6). A value of zero corresponds to a satellite which was out of view of the station at the time of retrieval; this response is not counted as either a detection or a non-detection. A value of one signifies a non-detection by a station in view of a satellite as it passes through the fence energy field. Any value of two or greater signifies a detection by the station; the detection intensity (amount of return power) increases with increasing numerical value. Only these cases are used as

detections while the others are considered as non-detections. By analyzing the results in Table 4.3, one can see that the percentages of detections for those stations located in the central portions of the radar fence are higher than those in the coastal regions. Notice that the highest percentage of detections are for receiving station five. This is not surprising since there is a greater number of satellites whose orbital paths fall within the longitudes associated with this area. The anomaly associated with the lower percentage of detections for station two may be due to the fact that this station was not yet fully operational at the time of data retrieval.

The six models are selected through the use of the stepwise logistic regression procedure. The backward elimination options were used at the significance level of 0.2, and the results are provided in Tables 4.4 and 4.5.

The parameter estimates given in Tables 4.4 and 4.5 are then used to the predict the probability of satellite detection when their associated characteristics $(x_{i1},..,x_{in})$ are given:

$$\hat{p}_i = \frac{\exp(\beta_o+\beta_1 * x_{i1}+\beta_2 * x_{i2}+....\beta_n * x_{in})}{1+\exp(\beta_c+\beta_1 * x_{i1}+\beta_2 * x_{i2}+....\beta_n * x_{in}} \qquad (4.1)$$

An example of how this equation is used to determine a station's prediction accuracy is provided later in the chapter.

2. Model Cross Validation

For the purpose of cross validation the classification (c-table) table is used. The c-table provides a measure of how robust the fitted model is regardless of changes

TABLE 4.4 PARAMETER ESTIMATES FOR INDIVIDUAL STATIONS

INDIVIDUAL STATION MODELS			
VARIABLES	STATION NUMBER/PARAMETER ESTIMATES		
	ONE	TWO	THREE
INTERCEPT (β_0)	85.5977	71.0796	52.9277
HOURS (β_1)	0.0513	0.0768	0.0610
RCS (β_2)	-0.0675	-0.0502	-0.0724
ECC (β_3)	2.4247	1.9915	2.0396
INCS (β_4)	0.5631	0.3999	0.5910
ALT (β_5)	-0.00272	-0.00227	-0.00200
ALT2 (β_6)	2.186E-8	1.948E-8	2.223E-8
LONG (β_7)	-0.2393	REMOVED	0.2033
LONG2 (β_8)	REMOVED	-0.00011	REMOVED
PER (β_9)	-0.00290	-0.00229	-0.00299
LAMBDA (β_{10})	-3.2039	-3.2293	-3.0347
SOLF (β_{11})	0.00089	-0.0127	REMOVED
LAMBDA2 (β_{13})	0.0246	0.0356	0.0393
LLMCROSS (β_{14})	0.00520	REMOVED	-0.00499
LACROSS (β_{15})	5.693E-6	9.972E-7	-2.06E-6
ALMCROSS (β_{16})	0.000061	0.000059	0.000064
A3CROSS (β_{17})	-4.03E-8	REMOVED	-4.29E-8
GEOM (β_{18})	-0.00460	-0.00878	REMOVED

TABLE 4.5 PARAMETER ESTIMATES FOR INDIVIDUAL MODELS

INDIVIDUAL STATION MODELS			
VARIABLES	STATION NUMBER/PARAMETER ESTIMATES		
	FOUR	FIVE	SIX
INTERCEPT (β^0)	55.7852	29.1509	40.2902
HOURS (β_1)	0.0604	0.0553	0.0540
RCS (β_2)	-0.0759	-0.0645	-0.0749
ECC (β_3)	3.5641	1.7262	2.8936
INCS (β_4)	0.6304	0.4697	0.7147
ALT (β_5)	-0.00218	-0.00151	-0.00187
ALT2 (β_6)	2.781E-8	2.511E-8	3.029E-8
LONG (β_7)	0.2515	0.3714	0.2629
LONG2 (β_8)	REMOVED	REMOVED	0.00114
PER (β_9)	-0.00448	-0.00237	-0.00345
LAMBDA (β_{10})	-3.3106	-2.1346	-2.6321
SOLF (β_{11})	-0.00078	0.00118	REMOVED
LAMBDA2 (β_{13})	0.0437	0.0304	0.0438
LLMCROSS (β_{14})	-0.00594	-0.00882	-0.0117
LACROSS (β_{15})	-4.43E-6	-7.73E-6	-8.42E-6
ALMCROSS (β_{16})	0.000072	0.000054	0.000069
A3CROSS (β_{17})	REMOVED	3.3825E-8	REMOVED
GEOM (β_{18})	-0.00380	-0.00758	-0.00460

in data. An example of the c-table output can be seen in Figure 4.1. The following values associated with the c-table results are defined for better understanding.

	Correct		Incorrect		Percentages				
Prob Level	Event	Non-Event	Event	Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.880	248	16E4	337	50891	75.7	0.5	99.8	57.6	24.2
0.900	204	16E4	271	50935	75.8	0.4	99.8	57.1	24.2
0.920	178	16E4	249	50961	75.7	0.3	99.8	58.3	24.2
0.940	143	16E4	219	50996	75.7	0.3	99.9	60.5	24.2
0.960	103	16E4	173	51036	75.7	0.2	99.9	62.7	24.2
0.980	59	16E4	124	51080	75.8	0.1	99.9	67.8	24.2
1.000	0	16E4	0	51139	75.8	0.0	100.0	.	24.2

Figure 4.1 C-Table Result Format

1. Probability Level: The level at which classifications are made. The latest SAS version provides classification results for each level starting at 0.0 through 1.00 at 0.02 increment [Ref. 9]. At each level the jackknife procedure outlined in Chapter Three is performed. Once the procedure is completed the predicted response is compared to the given probability level. If predicted response is greater than the given probability level then the response is classified as an event. If it is less than the probability level it is classified as a non-event. Event corresponds to response 0 (non-detection) while non-event corresponds to response 1 (detection) in the binary model case (used for individual station models). In a binomial model, one can specify the number of events (detection) versus the number of trials

(detection plus non-detection) in the model specification (used in system-wide model case).

2. Classification of predicted values: The next four lines of output provide the results of the classification [See Figure 4.1]. Each trial is classified in one of the four categories given in Table 4.6.

3. Correct: The percentage of predicted response either events or non-events that were correctly classified for the given probability level.

4. Sensitivity: The proportion of event responses that were predicted as events.

5. Specificivity: The proportion of non-event responses classified as non-events.

6. False positive rate: The proportion of predicted event responses that were observed as non-event responses.

7. False negative rate: The proportion of predicted non-events responses that were observed as events.

TABLE 4.6 CLASSIFICATION CATEGORIES

CATEGORY	OBSERVED	PREDICTED
1	EVENT	EVENT
2	EVENT	NON-EVENT
3	NON-EVENT	NON-EVENT
4	NON-EVENT	EVENT

By understanding the structure of the c-table one can select a probability level at which classifications can be made to minimize the sum of the two possible errors. In order to find the appropriate level for each station the minimum error sum rule is applied. This is done by first adding the sum of the false positive and false negative error rates at each probability level. The level which provides the lowest error sum is then selected as the appropriate probability level for classification. This is accomplished through the use of an additional SAS program [See Appendix A]. The probability levels selected for the six individual station models are provided in Table 4.7.

The following example is provided to demonstrate how the model and probability level selected for Station One can be used to demonstrate the station's detection performance capability. The values for the parameter estimates [See Table 4.4] and the values for variables ($x_1,..,x_{17}$) for SATCAT number 130 [See Table 4.8] are applied to equation 4.1 to determine the predicted probability ($\hat{p}$). The predicted

TABLE 4.7 PROBABILITY LEVELS FOR SIX STATIONS

STATION	LEVEL
ONE	0.36
TWO	0.34
THREE	0.32
FOUR	0.30
FIVE	0.32
SIX	0.34

probability ($\hat{p}$) for the values given is 0.34615. The value for $\hat{p}$ is now compared to the station's selected probability level. The value for $\hat{p}$ in this case is less than the probability level selected (0.36 for Station One) and is therefore classified as a non-event (detection). When checking the original raw data set, as a means of cross validation, it is seen that the actual response for this case was also classified as a non-event (detection). If it were greater than the selected probability level it would be classified as an event (non-detection).

C. SYSTEM-WIDE MODEL

This section consists of two subsections. Subsection one provides a description of the system-wide model that was generated using the SAS PROC LOGISTIC. Subsection two provides the cross validation results for the system-wide model. Additionally, subsection two provides a comparison between the model suggested in Schaaf [Ref. 1: p. 31] and the new system-wide model. The comparison is based upon the error sum

TABLE 4.8 VALUES FOR ORBITAL AND GEOMAGNETIC/SOLAR CHARACTERISTICS FOR SATCAT NUMBER 130

VARIABLE NAME	VALUE
HOURS (x_1)	1
RCS (x_2)	0.611
ECC (x_3)	0.0086
INCS (x_4)	1.16536
ALT (x_5)	844.20
ALT2 (x_6)	712673.64
LONG (x_7)	108.62
LONG2 (x_8)	REMOVED
PER (x_9)	103.1
LAMBDA (x_{10})	33.211
SOLF (x_{11})	135
LAMBDA2 (x_{12})	1102.97
LLMCROSS (x_{13})	3607.38
LACROSS (x_{14})	91697.00
ALMCROSS (x_{15})	28036.73
A3CROSS (x_{16})	3045349.20
GEOM (x_{17})	13

results obtained from the c-table.

1. Model Formulation

Based upon the same data set used for the individual station models, a prediction model for the entire system is estimated.

The response profile for the system-wide model is given in Table 4.9. The response profile for the system-wide model is based upon the total number of detections and non-detections for all six stations combined.

TABLE 4.9 RESPONSE PROFILE FOR SYSTEM-WIDE MODEL

FULL MODEL RESPONSE PROFILE	
EVENTS	TRIALS
51139	160026

The same set of independent variables used in the individual model selection was used in the full model exploration. No variables met the criterion for removal. The final model for the prediction of the system-wide performance is given in Table 4.10. Using equation 4.1, one can obtain the predicted probability of detection for the given characteristics of the satellite. A complete printout of the output provided by SAS is provided in Appendix A.

2. Cross Validation and Comparisons

In this section the optimum probability level for each model is selected from a comparison of error sums provided by the c-table. Based on these results one can compare the prediction accuracy of the two models. The results are given in Table 4.11. It can be seen in Table 4.11 that the lowest error sum for the new model occurred at a probability level of 0.60 and the lowest error sum for the old model occurred at 0.48. When comparing the two error sums, it is observed that the error sum for the new model (52) is lower than that of the old model (59.5). This indicates that the new model is

TABLE 4.10 SYSTEM-WIDE MODEL PARAMETER ESTIMATES

SYSTEM-WIDE MODEL	
VARIABLE	PARAMETER ESTIMATE
INTERCEPT (β_0)	-23.0699
HOURS (β_1)	-0.0553
RCS (β_2)	0.0599
ECC (β_3)	-2.0058
INCS (β_4)	-0.5644
ALT (β_5)	0.00105
ALT2 (β_6)	-1.57E-8
LONG (β_7)	-0.2402
LONG2 (β_8)	0.00116
PER (β_9)	0.00268
LAMBDA (β_{10})	1.4086
SOLF (β_{11})	0.00132
LAMBDA2 (β_{12})	-0.00853
LLMCROSS (β_{13})	-0.0005
LACROSS (β_{14})	-8.06E-7
ALMCROSS (β_{15})	-0.00003
A3CROSS (β_{16})	6.272E-8
GEOM (β_{17})	0.00510

better at properly classifying detections and non-detections than the old model. At these selected probability levels classification can be made and the results are compared in Table 4.11. These values are used to establish a probability level for classification in the same manner as the individual station models.

TABLE 4.11 SYSTEM-WIDE C-TABLE RESULTS

COMBINED C-TABLE RESULTS			
MODEL	LEVEL	CORRECT EVENTS	CORRECT NON-EVENTS
NEW	0.60	3621	159000
OLD	0.48	5578	157000
MODEL	INCORRECT EVENTS	INCORRECT NON-EVENTS	ERROR SUM
NEW	1471	47518	52.0
OLD	3275	45561	59.5

The following example is provided to demonstrate how the probability level selected for the system-wide case can be used to demonstrate the whole system's prediction performance capability. The values for the parameter estimates [See Table 4.10] and the variables (x_1,..,x_{17}) for SATCAT number 63 [Table 4.12] are applied to equation 4.1 to determine the predicted probability ($\hat{p}$). The calculated value of the predicted response ($\hat{p}$) for the values given is 0.60426. The value for $\hat{p}$ is now compared to the selected probability level (0.60). For this case $\hat{p}$ is greater than the probability level established for the system-wide model and is therefore classified as a detection. When comparing this result to the actual raw data set responses, as a means of cross validation, it is seen that every station in this case recorded a detection for this particular SATCAT case.

TABLE 4.12 VALUES FOR ORBITAL AND GEOMAGNETIC/SOLAR CHARACTERISTICS FOR SATCAT 63

SYSTEM-WIDE MODEL EXAMPLE PROBLEM VARIABLES	
VARIABLE NAME	VALUE
HOURS (x_1)	3
RCS (x_2)	.357
ECC (x_3)	0.0049
INCS (x_4)	0.84683
ALT (x_5)	596.66
ALT2 (x_6)	356003.16
LONG (x_7)	97.38
LONG2 (x_8)	9289.10
PER (x_9)	96.30
LAMBDA (x_{10})	33.323
SOLF (x_{11})	135
LAMBDA2 (x_{12})	1110.42
LLMCROSS (x_{13})	3211.67
LACROSS (x_{14})	57506.09
ALMCROSS (x_{15})	19882.50
A3CROSS (x_{16})	1916275.46
GEOM (x_{17})	13

V. CONCLUSIONS

The objective of this research was to provide an improved prediction model for measuring the NAVSPASUR radar fence performance. In doing so six additional individual station models and a new improved system-wide model have been provided. Additionally, this research has provided probability levels for each of the seven models, which were not previously provided. These values are used to establish levels for classifying the detection capability of the system at the minimum level of error. The six additional individual station models provide NAVSPASUR with an additional capability to assess individual performance that, up to this point was not available. The new system-wide model, still not fully explored, is superior to the one previously provided in terms of its accuracy of classification. These results are evident in the comparisons made between the two model's c-tables provided in Chapter IV. The increase in prediction accuracy may be due to the use of additional variables not previously analyzed. The solar and geomagnetic variables added, though not extremely influential in some cases, did in most cases add to the overall system and individual system prediction accuracies. Also, it is apparent from the background provided in Chapter II that any variable that affects the range between an object and the radar itself has a great deal of influence upon the radar detection capability. All these variables have proven to be statistically significant when they are used together for estimating the prediction models provided. The results of this research have provided a valuable tool

(FORTRAN Implementation Program) that allows NAVSPASUR operators to test their system's detection capability statistically, at any time.

Recommendations for possible areas of further research follow. Possible seasonal effects on system performance could be analyzed by using additional data (preferably some portion from of each month over the span of a full year). One could possibly show variations that may be introduced by seasonal effects. An attempt to integrate more explanatory variables than those analyzed here could also increase the accuracy of the prediction model. Using new statistical analysis techniques such as probit, complementary log-log or random-effect logistic regression analysis could also possibly provide new insight.

APPENDIX A

A. FORTRAN CODING AND RELATED OUTPUT

1. Data Set Manipulation Programs

The following two programs written in FORTRAN code are used to generate the data set used in the logistic regression analysis.

1. Sampling Program. This program was written to select a random sample of data from the original data sets provided by NAVSPASUR.

```
//SAMPLE   JOB (8088,9999),'SAMPLE FORTRAN',CLASS=B
// EXEC VSF2CG,IMSL=IMSL10
//FORT.SYSIN DD *
      INTEGER SATCAT
      CHARACTER*80 DATA(300)
      CHARACTER*11 DATA2(300)
C      ISEED MUST BE IN THE RANGE OF (0,2147483646)
      ISEED = 346789123
      CALL RNSET(ISEED)
      READ (1,10,END=100) SATCAT,DATA(1),DATA2(1)
      OLDCAT=SATCAT
       NCAT = 2
5      READ (1,10,END=100) SATCAT,DATA(NCAT),DATA2(NCAT)
10     FORMAT (T18,I6,T1,A80,/,A11)
      IF (OLDCAT .EQ. SATCAT ) THEN
      NCAT= NCAT + 1
      ELSE
      CALL OUTPUT(NCAT,DATA,DATA2)
      DATA(1) = DATA(NCAT+1)
      DATA2(1) = DATA2(NCAT+1)
      OLDCAT=SATCAT
      NCAT = 2
      END IF
      GO TO 5
```

```
100  CONTINUE
     END
     SUBROUTINE OUTPUT(NCAT,DATA,DATA2)
     CHARACTER*80 DATA(300)
     CHARACTER*11 DATA2(300)
     DIMENSION R(300)
     REAL LRN
     NCAT= NCAT - 1
     CALL RNUN (NCAT,R)
   LRN   = 0.0
     LINDEX = 0.0
     DO 20 I = 1,NCAT
   IF (LRN .LT. R(I)) THEN
     LINDEX = I
     LRN = R(I)
     END IF
20   CONTINUE
     WRITE(2,10) DATA(LINDEX),DATA2(LINDEX)
10   FORMAT(A80,A11)
     RETURN
     END
/*
//GO.FT01F001 DD DISP=SHR,DSN=MSS.S8088.SATM07.DATA
//*
//* IF THE PROGRAM MUST BE RUN AGAIN WITH THE SAME DATA
//* SET, CHANGE NEW,CATLG TO OLD,KEEP
//*
//GO.FT02F001 DD UNIT=SYSDA,DISP=(OLD,KEEP),
//   DCB=(RECFM=FB,LRECL=91,BLKSIZE=23387),
//   SPACE=(23387,(24,3)),
//   DSNAME=MSS.S8088.SATM07.DATA.SINGLE
//
//THESFUV  JOB (8088,9999),'THESFUV  SAS',CLASS=C
// EXEC SAS,REGION=7872K
//WORK DD UNIT=SYSDA,SPACE=(CYL,(16,16))
//SASIN DD DISP=SHR,DSN=MSS.S8088.FINAL
//SYSIN DD *
OPTIONS LS=80;
PROC UNIVARIATE DATA=SASIN.FINAL; VAR RCS ECC INC ALT LONG
PER LAMBDA
SOLF ASOLF GEOM GEOMDAY;
```

2. Data set merge program. This program merges the data

set created by the sampling program with the proper solar/geomagnetic data provided by NAVSPASUR.

```
//MERGE    JOB (8088,9999),'MERGE FORTRAN',CLASS=B
// EXEC VSF2CG,IMSL=IMSL10
//FORT.SYSIN DD *
      CHARACTER * 6 SDATE(10),DDATE
      CHARACTER * 19 SDATA(10)
      CHARACTER * 3 SFLUX(10,8)
      CHARACTER *83 DDATA
      INTEGER DHOUR
      I=1
5     READ (1,10,END=20)
      SDATE(I),SDATA(I),(SFLUX(I,J),J=1,8)
10    FORMAT(A6,A19,8(4X,A3))
      I=I+1
      GOTO 5
20    NSOL=I-1
      DO 100 I=1,NSOL
      DO 100 J=1,8
100   WRITE (6,110) SDATE(I),SDATA(I),SFLUX(I,J)
110   FORMAT(1X,A6,1X,A19,1X,A3)
210   READ(2,200,END=300) DDATE,DHOUR,DDATA
200   FORMAT(1X,A6,I2,A82)
      J=(DHOUR+3)/3
      DO 250 I=1,NSOL
      IF (DDATE .EQ. SDATE(I)) GO TO 270
250   CONTINUE
      WRITE(6,260) DDATE
260   FORMAT(1X,A6,' DATE DOES NOT MATCH SOLAR DATA')
      GO TO 210
270   WRITE(3,280) DDATE,DHOUR,DDATA,SDATA(I),SFLUX(I,J)
280   FORMAT(A6,I2,A82,1X,A19,8(1X,A3))
      GO TO 210
300   CONTINUE
      END
/*
//GO.FT01F001 DD DISP=SHR,DSN=MSS.S8088.SOL0411.DATA
//GO.FT02F001 DD DISP=SHR,DSN=MSS.S8088.SATM05.DATA
//*
//*  IF THE PROGRAM MUST BE RUN AGAIN WITH THE SAME DATA
//*  SET CHANGE NEW,CATLG TO OLD,KEEP
//*
//GO.FT03F001 DD UNIT=SYSDA,DISP=(OLD,KEEP),
//    DCB=(RECFM=FB,LRECL=114,BLKSIZE=23370),
//    SPACE=(23370,(24,3)),
//    DSNAME=MSS.S8088.SATM05.DATA.MFULL
//
```

2. Fortran Implementation Program

The program when compiled in a standard compiler creates a file that can then be executed on any IBM compatible computer. The program will prompt the user to enter the variables to be analyzed. Once this is done the user will be asked which station (stations one through six or system-wide) he/she would like to analyze. At that point the program computes the predicted probability of detection for the case entered. The predicted probability is then compared to the probability level assigned for the particular station and classifies the prediction as either a detection or a non-detection. At this point the program gives the user three options; the user can the either enter new variables, select a new station, or exit the program.

```
      PROGRAM SPACUR
C    SPACUR IS A PROGRAM THAT ALLOWS RADAR FENCE OPERATORS TO
C     MEASURE SYSTEM PERFORMANCE FOR A GIVEN SET OF ORBITAL,
C     GEOMAGNETIC AND SOLAR CHARACTERISTICS.
C
C     DECLARE INTEGERS
C     LOOP IS AN INPUT WITCH TO DECIDE WHETHER THE OPERATOR
C     DESIRES TO ENTER NEW PARAMETERS.
      INTEGER LOOP
C    LOOP1 IS AN INPUT SWITCH TO DECIDE WHETHER THE OPERATOR
C     DESIRES TO ENTER A NEW STATION NUMBER OR EXIT THE
C     PROGRAM.
      INTEGER LOOP1
C
C     DECLARE REALS
C
      REAL PH1,PH2,PH3,PH4,PH5,PH6,PH7
      REAL PVAL1,PVAL2,PVAL3,PVAL4,PVAL5,PVAL6,PVAL7
      REAL HOUR
      REAL RCS
```

```
      REAL INC
      REAL INCS
      REAL ALT
      REAL LONG
      REAL PER
      REAL LAT
      REAL SOLF
      REAL ASOLF
      REAL GEOM
      REAL GEOMD
      REAL ECC
      REAL ALT2
      REAL LONG2
      REAL LAT2
      REAL LLMC
      REAL LAMC
      REAL LALT
      REAL A3C
      REAL A1,A2,A3,A4,A5,A6,A7
      REAL B1,B2,B3,B4,B5,B6,B7
      REAL C1,C2,C3,C4,C5,C6,C7
      REAL D1,D2,D3,D4,D5,D6,D7
      REAL E1,E2,E3,E4,E5,E6,E7
      REAL F1,F2,F3,F4,F5,F6,F7
      REAL G1,G2,G3,G4,G5,G6,G7
      REAL H1,H2,H3,H4,H5,H6,H7
      REAL I1,I2,I3,I4,I5,I6,I7
      REAL J1,J2,J3,J4,J5,J6,J7
      REAL K1,K2,K3,K4,K5,K6,K7
      REAL L1,L2,L3,L4,L5,L6,L7
      REAL M1,M2,M3,M4,M5,M6,M7
      REAL N1,N2,N3,N4,N5,N6,N7
      REAL O1,O2,O3,O4,O5,O6,O7
      REAL P1,P4,P5,P6,P7
      REAL Q1,Q5
      REAL P1A,P2A,P3A,P4A,P5A,P6A,P7A
      REAL R7
C
C     INPUT THE PARAMETERS.
C
200   CONTINUE
      WRITE(6,*)'INPUT HOUR'
      READ(5,*)HOUR
      WRITE(6,*)'HOUR=',HOUR
      WRITE(6,*)'INPUT RADAR CROSS SECTION'
      READ(5,*)RCS
      WRITE(6,*)'RCS=',RCS
      WRITE(6,*)'INPUT ECCENTRICITY'
      READ(5,*)ECC
      WRITE(6,*)'ECC=',ECC
```

```
      WRITE(6,*)'INPUT INCLINATION'
      READ(5,*)INC
      WRITE(6,*)'INC=',INC
      WRITE(6,*)'INPUT ALTITUDE'
      READ(5,*)ALT
      WRITE(6,*)'ALT=',ALT
      WRITE(6,*)'INPUT LONGITUDE'
      READ(5,*)LONG
      WRITE(6,*)'LONG=',LONG
      WRITE(6,*)'INPUT PERIOD'
      READ(5,*)PER
      WRITE(6,*)'PER=',PER
      WRITE(6,*)'INPUT LATITUDE'
      READ(5,*)LAT
      WRITE(6,*)'LAT=',LAT
      WRITE(6,*)'INPUT SOLAR FLUX'
      READ(5,*)SOLF
      WRITE(6,*)'SOLF=',SOLF
      WRITE(6,*)'INPUT AVERAGE SOLAR FLUX'
      READ(5,*)ASOLF
      WRITE(6,*)'ASOL=',ASOLF
      WRITE(6,*)'INPUT GEOMAGNETISM'
      READ(5,*)GEOM
      WRITE(6,*)'GEOM=',GEOM
      WRITE(6,*)'INPUT DAILY AVERAGE'
      READ(5,*)GEOMD
      WRITE(6,*)'GEOMD=',GEOMD
C
C     DETERMINE HIGHER ORDER TERMS
C
      INCS=ACOS(INC)
      ALT2=ALT*ALT
      LONG2=LONG*LONG
      LAT2=LAT*LAT
      LLMC=LAT*LONG
      LALT=LONG*ALT
      LAMC=LAT*ALT
      A3C=ALT*LAT*LONG
C
C     SELECT THE STATION NUMBER DESIRED.
C
500   WRITE(6,*)'INPUT SELECTED STATION NUMBER:'
      WRITE(6,*)'1=SAN DIEGO RECEIVER'
      WRITE(6,*)'2=ELEPHANT BUTTE RECEIVER'
      WRITE(6,*)'3=RED RIVER RECEIVER'
      WRITE(6,*)'4=SILVER LAKE RECEIVER'
      WRITE(6,*)'5=HAWKINSVILLE RECEIVER'
      WRITE(6,*)'6=TATTNAL RECEIVER'
      WRITE(6,*)'7=GLOBAL MODEL'
      READ(5,*)STATN
```

```
      IF(STATN/(INT(STATN)) .NE. 1.0) THEN
      WRITE(6,*)'STATION NUMBER MUST BE REAL INTEGER 1-7!!'
      GOTO 500
      ELSE
      ENDIF
      IF (STATN .LT. 0.999) THEN
      WRITE(6,*)'STATION NUMBER MUST BE REAL INTEGER 1-7!!'
      GOTO 500
      ELSE
      ENDIF
      IF (STATN .GT. 7.001) THEN
      WRITE(6,*)'STATION NUMBER MUST BE REAL INTEGER 1-7!!'
      GOTO 500
      ELSE
      ENDIF
      IF(STATN .EQ. 1) THEN
      GOTO 1000
      ELSEIF (STATN .EQ. 2) THEN
      GOTO 2000
      ELSEIF (STATN .EQ. 3) THEN
      GOTO 3000
      ELSEIF (STATN .EQ. 4) THEN
      GOTO 4000
      ELSEIF (STATN .EQ. 5) THEN
      GOTO 5000
      ELSEIF (STATN .EQ. 6) THEN
      GOTO 6000
      ELSEIF (STATN .EQ. 7) THEN
      GOTO 7000
      ELSE
      ENDIF
C
C     PERFORM THE STATION 1 CALCULATIONS.
C
1000  CONTINUE
      A1=85.5977
      B1=0.0513*HOUR
      C1=-0.0675*RCS
      D1=2.4247*ECC
      E1=.5631*INCS
      F1=-0.00272*ALT
      G1=2.186E-8*ALT2
      H1=-.2392*LONG
      I1=-0.00290*PER
      J1=-3.2039*LAT
      K1=.00089*SOLF
      L1=.0246*LAT2
      M1=.00520*LLMC
      N1=5.693E-6*LALT
      O1=.000061*LAMC
```

```
      P1=-4.03E-8*A3C
      Q1=-0.00460*GEOMD
C
      PH1=A1+B1+C1+D1+E1+F1+G1+H1+I1+J1+K1+L1+M1+N1+O1+P1+Q1
      WRITE(6,*)'PH1=',PH1
      P1A=EXP(PH1)
      WRITE(6,*)P1A
      PVAL1=P1A/(1+P1A)
      WRITE(6,*)'PVALUE=',PVAL1
      IF (PVAL1 .LE. 0.64) THEN
      WRITE(6,*)'VALID DETECTION'
      ELSE
      WRITE(6,*)'INVALID DETECTION'
      ENDIF
      GOTO 8000
2000  CONTINUE
      A2=71.0796
      B2=0.0768*HOUR
      C2=-0.0502*RCS
      D2=1.9915*ECC
      E2=.3999*INCS
      F2=-0.00227*ALT
      G2=1.948E-8*ALT2
      H2=-0.00011*LONG2
      I2=-0.00229*PER
      J2=-3.2293*LAT
      K2=-0.0127*SOLF
      L2=0.0356*LAT2
      M2=9.927E-7*LALT
      N2=0.000059*LAMC
      O2=-0.00878*GEOMD
C
      PH2=A2+B2+C2+D2+E2+F2+G2+H2+I2+J2+K2+L2+M2+N2+O2
      WRITE(6,*)'PH2=',PH2
      P2A=EXP(PH2)
      WRITE(6,*)P2A
      PVAL2=P2A/(1+P2A)
      WRITE(6,*)'PVALUE=',PVAL2
      IF (PVAL2 .LE. 0.66) THEN
      WRITE(6,*)'VALID DETECTION'
      ELSE
      WRITE(6,*)'INVALID DETECTION'
      ENDIF
      GOTO 8000
3000  CONTINUE
      A3=52.9277
      B3=0.0610*HOUR
      C3=-0.0724*RCS
      D3=2.0396*ECC
      E3=0.591*INCS
```

```
      F3=-0.002*ALT
      G3=2.223E-8*ALT2
      H3=0.2033*LONG
      I3=-0.00299*PER
      J3=-3.0347*LAT
      K3=0.0393*LAT2
      L3=-0.00499*LLMC
      M3=-2.06E-6*LALT
      N3=0.000064*LAMC
      O3=-4.28E-8*LAMC
C
      PH3=A3+B3+C3+D3+E3+F3+G3+H3+I3+J3+K3+L3+M3+N3+O3
      WRITE(6,*)'PH3=',PH3
      P3A=EXP(PH3)
      WRITE(6,*)P3A
      PVAL3=P3A/(1+P3A)
      WRITE(6,*)'PVALUE=',PVAL3
      IF (PVAL3 .LE. 0.68) THEN
      WRITE(6,*)'VALID DETECTION'
      ELSE
      WRITE(6,*)'INVALID DETECTION'
      ENDIF
      GOTO 8000
4000  CONTINUE
      A4=55.7852
      B4=0.0604*HOUR
      C4=-0.0759*RCS
      D4=3.5641*ECC
      E4=0.6304*INCS
      F4=-0.00218*ALT
      G4=2.781E-8*ALT2
      H4=0.2515*LONG
      I4=-0.00488*PER
      J4=-3.3106*LAT
      K4=-0.00078*SOLF
      L4=0.0437*LAT2
      M4=-0.00594*LLMC
      N4=-4.43E-6*LALT
      O4=.000072*LAMC
      P4=-0.00380*GEOMD
C
      PH4=A4+B4+C4+D4+E4+F4+G4+H4+I4+J4+K4+L4+M4+N4+O4+P4
      WRITE(6,*)'PH4=',PH4
      P4A=EXP(PH4)
      WRITE(6,*)P4A
      PVAL4=P4A/(1+P4A)
      WRITE(6,*)'PVALUE=',PVAL4
      IF (PVAL4 .LE. 0.70) THEN
      WRITE(6,*)'VALID DETECTION'
      ELSE
```

```
      WRITE(6,*)'INVALID DETECTION'
      ENDIF
      GOTO 8000
 5000 CONTINUE
      A5=29.1509
      B5=0.0553*HOUR
      C5=-0.0645*RCS
      D5=1.7262*ECC
      E5=0.4697*INCS
      F5=-0.00151*ALT
      G5=2.511E-8*ALT2
      H5=0.3714*LONG
      I5=-0.00237*PER
      J5=-2.1346*LAT
      K5=0.00118*SOLF
      L5=0.0304*LAT2
      M5=-0.00882*LLMC
      N5=-7.73E-6*LALT
      O5=0.000054*LAMC
      P5=3.382E-8*A3C
      Q5=-0.00758*GEOMD
C
      PH5=A5+B5+C5+D5+E5+F5+G5+H5+I5+J5+K5+L5+M5+N5+O5+P5+Q5
      WRITE(6,*)'PH5=',PH5
      P5A=EXP(PH5)
      WRITE(6,*)P5A
      PVAL5=P5A/(1+P5A)
      WRITE(6,*)'PVALUE=',PVAL5
      IF (PVAL5 .LE. 0.68) THEN
      WRITE(6,*)'VALID DETECTION'
      ELSE
      WRITE(6,*)'INVALID DETECTION'
      ENDIF
      GOTO 8000
 6000 CONTINUE
      A6=4.2902
      B6=0.0540*HOUR
      C6=-0.0749*RCS
      D6=2.8396*ECC
      E6=0.7147*INCS
      F6=-0.00187*ALT
      G6=3.029E-8*ALT2
      H6=0.2629*LONG
      I6=0.00114*LONG2
      J6=-0.00345*PER
      K6=-2.321*LAT
      L6=0.0438*LAT2
      M6=-0.0117*LLMC
      N6=-8.42E-6*LALT
      O6=0.000069*LAMC
```

```
      P6=-0.0046*GEOMD
C
      PH6=A6+B6+C6+D6+E6+F6+G6+H6+I6+J6+K6+L6+M6+N6+O6+P6
      WRITE(6,*)'PH6=',PH6
      P6A=EXP(PH6)
      WRITE(6,*)P6A
      PVAL6=P6A/(1+P6A)
      WRITE(6,*)'PVALUE=',PVAL6
      IF (PVAL6 .LE. 0.66) THEN
      WRITE(6,*)'VALID DETECTION'
      ELSE
      WRITE(6,*)'INVALID DETECTION'
      ENDIF
      GOTO 8000
7000  CONTINUE
      A7=-23.0699
      B7=-0.0553*HOUR
      C7=0.0599*RCS
      D7=-2.0058*ECC
      E7=-0.5644*INCS
      F7=0.00105*ALT
      G7=-1.57E-8*ALT2
      H7=-0.2402*LONG
      I7=0.00116*LONG2
      J7=0.00268*PER
      K7=1.4086*LAT
      L7=0.00132*SOLF
      M7=-0.00853*LAT2
      N7=-0.0005*LLMC
      O7=-8.06E-7*LALT
      P7=-0.00003*LAMC
      Q7=6.272E-8*A3C
      R7=0.00510*GEOMD
C

PH7=A7+B7+C7+D7+E7+F7+G7+H7+I7+J7+K7+L7+M7+N7+O7+P7+Q7+R7
      WRITE(6,*)'PH7=',PH7
      P7A=EXP(PH7)
      WRITE(6,*)P7A
      PVAL7=P7A/(1+P7A)
      WRITE(6,*)'PVALUE=',PVAL7
      IF (PVAL7 .GE. 0.40) THEN
      WRITE(6,*)'VALID DETECTION'
      ELSE
      WRITE(6,*)'INVALID DETECTION'
      ENDIF
      GOTO 8000
C
C     SELECT WHETHER TO ENTER NEW PARAMETERS.
C
```

```
8000    WRITE(6,*)'DO  YOU  WISH  TO  MANUALLY  ENTER  NEW
PARAMETERS?'
        WRITE(6,*)'1=YES'
        WRITE(6,*)'2=NO'
        READ(5,*)LOOP
        IF (LOOP .EQ. 1) THEN
        GOTO 200
        ELSE
        ENDIF
C
C        SELECT WHETHER TO SELECT A NEW STATION OR END THIS
SESSION.
C
        WRITE(6,*)'DO YOU WISH TO SELECT A NEW STATION OR '
        WRITE(6,*)'EXIT THE PROGRAM?'
        WRITE(6,*)'1=SELECT NEW STATION'
        WRITE(6,*)'2=EXIT PROGRAM'
        READ(5,*)LOOP1
        IF(LOOP1 .EQ. 1) THEN
        GOTO 500
        ELSE
        ENDIF
C
        END
```

B. SAS PROGRAMMING AND RELATED OUTPUT

In this section the SAS programs and related output referenced in Chapters III and IV are provided.

1. Univariate program

```
//THESFUV  JOB (8088,9999),'THESFUV  SAS',CLASS=C
//   EXEC SAS,REGION=7872K
//WORK DD UNIT=SYSDA,SPACE=(CYL,(16,16))
//SASIN DD DISP=SHR,DSN=MSS.S8088.FINAL
//SYSIN DD *
OPTIONS LS=80;
PROC UNIVARIATE DATA=SASIN.FINAL; VAR RCS ECC INC ALT LONG PER LAMBDA
SOLF ASOLF GEOM GEOMDAY;
```

2. Station One Program and Output.

```
        OPTIONS LS=80;
        PROC LOGISTIC DATA=SASIN.FINAL;
        MODEL Y1=HOURS RCS ECC INCS ALT ALT2 LONG LONG2 PER LAMBDA SOLF
        LAMBDA2 LLMCROSS LACROSS ALMCROSS A3CROSS
        GEOMDAY / CTABLE SELECTION=B SLSTAY=.2 FAST;

                              THE SAS SYSTEM

                          The LOGISTIC Procedure

   Data Set: SASIN.FINAL
   Response Variable: Y1
```

Response Levels: 2
Number of Observations: 47464
Link Function: Logit

Response Profile

Ordered Value	Y1	Count
1	0	39037
2	1	8427

Backward Elimination Procedure

Step 0. The following variables were entered:

INTERCPT	HOURS	RCS	ECC	INCS	ALT	ALT2
LONG	LONG2	PER	LAMBDA	SOLF	LAMBDA2	LLMCROSS
LACROSS	ALMCROSS	A3CR S	GEOMDAY			

Criteria for Assessing Model Fit

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	44395.126	29940.938	.
SC	44403.894	30098.757	.
-2 LOG L	44393.126	29904.938	14488.189 with 17 DF (p=0.0001)
Score	.	.	9351.898 with 17 DF (p=0.0001)

Step 1. Fast Backward Elimination:

Analysis of Variables Removed by Fast Backward Elimination

Variable Removed	Chi-Square	Pr > Chi-Square	Residual Chi-Square	DF	Pr > Residual Chi-Square
LONG2	0.0312	0.8598	0.0312	1	0.8598

Criteria for Assessing Model Fit

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	44395.126	29938.969	.
SC	44403.894	30088.020	.
-2 LOG L	44393.126	29904.969	14488.158 with 16 DF (p=0.0001)
Score	.	.	9351.422 with 16 DF (p=0.0001)

Residual Chi-Square = 0.0312 with 1 DF (p=0.8598)

The LOGISTIC Procedure

Summary of Backward Elimination Procedure

Step	Variable Removed	Number In	Wald Chi-Square	Pr > Chi-Square
1	LONG2	16	0.0312	0.8598

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	85.5977	6.1781	191.9609	0.0001	.	999.000
HOURS	1	0.0513	0.00219	550.7006	0.0001	0.190547	1.053
RCS	1	-0.0675	0.00242	777.6846	0.0001	-0.307552	0.935
ECC	1	2.4247	0.1906	161.8498	0.0001	0.272103	11.299
INCS	1	0.5631	0.0519	117.6954	0.0001	0.113155	1.756
ALT	1	-0.00272	0.00015	330.0916	0.0001	-9.303767	0.997
ALT2	1	2.186E-8	9.56E-10	522.6147	0.0001	2.321996	1.000
LONG	1	-0.2393	0.0282	72.0542	0.0001	-4.150754	0.787
PER	1	-0.00290	0.000302	92.3332	0.0001	-0.333189	0.997
LAMBDA	1	-3.2039	0.2901	122.0175	0.0001	-3.807776	0.041
SOLF	1	0.00089	0.000531	2.8115	0.0936	0.014775	1.001
LAMBDA2	1	0.0246	0.00335	53.9578	0.0001	3.205047	1.025
LLMCROSS	1	0.00520	0.000866	35.9708	0.0001	2.736208	1.005
LACROSS	1	5.693E-6	8.185E-7	48.3806	0.0001	2.190222	1.000
ALMCROSS	1	0.000061	4.254E-6	202.9887	0.0001	4.803480	1.000
A3CROSS	1	-4.03E-8	2.488E-8	2.6290	0.1049	-0.345834	1.000
GEOMDAY	1	-0.00460	0.00275	2.8017	0.0942	-0.014418	0.995

Association of Predicted Probabilities and Observed Responses

Concordant	= 87.8%	Somers' D	= 0.758
Discordant	= 12.0%	Gamma	= 0.760
Tied	= 0.2%	Tau-a	= 0.222
(328964799 pairs)		c	= 0.879

Classification Table

Prob Level	Correct Event	Correct Non-Event	Incorrect Event	Incorrect Non-Event	Percentages Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.000	39037	0	8427	0	82.2	100.0	0.0	17.8	.
0.020	39014	16	8411	23	82.2	99.9	0.2	17.7	59.0
0.040	39006	22	8405	31	82.2	99.9	0.3	17.7	58.5
0.060	38994	26	8401	43	82.2	99.9	0.3	17.7	62.3
0.080	38987	37	8390	50	82.2	99.9	0.4	17.7	57.5
0.100	38977	41	8386	60	82.2	99.8	0.5	17.7	59.4
0.120	38975	46	8381	62	82.2	99.8	0.5	17.7	57.4
0.140	38969	58	8369	68	82.2	99.8	0.7	17.7	54.0
0.160	38964	72	8355	73	82.2	99.8	0.9	17.7	50.3
0.180	38959	89	8338	78	82.3	99.8	1.1	17.6	46.7
0.200	38953	113	8314	84	82.3	99.8	1.3	17.6	42.6
0.220	38941	147	8280	96	82.4	99.8	1.7	17.5	39.5
0.240	38928	185	8242	109	82.4	99.7	2.2	17.5	37.1
0.260	38907	241	8186	130	82.5	99.7	2.9	17.4	35.0
0.280	38895	298	8129	142	82.6	99.6	3.5	17.3	32.3
0.300	38874	348	8079	163	82.6	99.6	4.1	17.2	31.9
0.320	38842	431	7996	195	82.7	99.5	5.1	17.1	31.2
0.340	38808	563	7864	229	82.9	99.4	6.7	16.8	28.9
0.360	38746	747	7680	291	83.2	99.3	8.9	16.5	28.0
0.380	38648	941	7486	389	83.4	99.0	11.2	16.2	29.2
0.400	38534	1167	7260	503	83.6	98.7	13.8	15.9	30.1
0.420	38361	1427	7000	676	83.8	98.3	16.9	15.4	32.1
0.440	38152	1706	6721	885	84.0	97.7	20.2	15.0	34.2
0.460	37911	2045	6382	1126	84.2	97.1	24.3	14.4	35.5
0.480	37629	2393	6034	1408	84.3	96.4	28.4	13.8	37.0
0.500	37309	2757	5670	1728	84.4	95.6	32.7	13.2	38.5

Prob Level	Correct Event	Correct Non-Event	Incorrect Event	Incorrect Non-Event	Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.520	36963	3166	5261	2074	84.5	94.7	37.6	12.5	39.6
0.540	36585	3561	4866	2452	84.6	93.7	42.3	11.7	40.8
0.560	36117	3925	4502	2920	84.4	92.5	46.6	11.1	42.7
0.580	35622	4324	4103	3415	84.2	91.3	51.3	10.3	44.1
0.600	35072	4709	3718	3965	83.8	89.8	55.9	9.6	45.7
0.620	34481	5070	3357	4556	83.3	88.3	60.2	8.9	47.3
0.640	33850	5398	3029	5187	82.7	86.7	64.1	8.2	49.0
0.660	33211	5778	2649	5826	82.1	85.1	68.6	7.4	50.2
0.680	32611	6077	2350	6426	81.5	83.5	72.1	6.7	51.4
0.700	31932	6336	2091	7105	80.6	81.8	75.2	6.1	52.9
0.720	31277	6628	1799	7760	79.9	80.1	78.7	5.4	53.9
0.740	30669	6846	1581	8368	79.0	78.6	81.2	4.9	55.0
0.760	30070	7050	1377	8967	78.2	77.0	83.7	4.4	56.0
0.780	29415	7237	1190	9622	77.2	75.4	85.9	3.9	57.1
0.800	28720	7412	1015	10317	76.1	73.6	88.0	3.4	58.2
0.820	28021	7565	862	11016	75.0	71.8	89.8	3.0	59.3
0.840	27296	7746	681	11741	73.8	69.9	91.9	2.4	60.3
0.860	26505	7862	565	12532	72.4	67.9	93.3	2.1	61.4
0.880	25606	7965	462	13431	70.7	65.6	94.5	1.8	62.8
0.900	24624	8077	350	14413	68.9	63.1	95.8	1.4	64.1
0.920	23504	8153	274	15533	66.7	60.2	96.7	1.2	65.6
0.940	22172	8235	192	16865	64.1	56.8	97.7	0.9	67.2
0.960	20569	8304	123	18468	60.8	52.7	98.5	0.6	69.0
0.980	18085	8356	71	20952	55.7	46.3	99.2	0.4	71.5
1.000	149	8427	0	38888	18.1	0.4	100.0	0.0	82.2

2. Error Sum Program

```
*STATION ONE FULL*;
OPTIONS LS=80;
DATA THRESH1F;
INPUT LEVEL E1 E2 E3 E4 CORRECT SEN SPEC FPOS FNEG;
CARDS;
      0.000  39037     0   8427     0    82.2  100.0    0.0  17.8     .
      0.020  39014    16   8411    23    82.2   99.9    0.2  17.7  59.0
      0.040  39006    22   8405    31    82.2   99.9    0.3  17.7  58.5
      0.060  38994    26   8401    43    82.2   99.9    0.3  17.7  62.3
      0.080  38987    37   8390    50    82.2   99.9    0.4  17.7  57.5
      0.100  38977    41   8386    60    82.2   99.8    0.5  17.7  59.4
      0.120  38975    46   8381    62    82.2   99.8    0.5  17.7  57.4
      0.140  38969    58   8369    68    82.2   99.8    0.7  17.7  54.0
      0.160  38964    72   8355    73    82.2   99.8    0.9  17.7  50.3
      0.180  38959    89   8338    78    82.3   99.8    1.1  17.6  46.7
      0.200  38953   113   8314    84    82.3   99.8    1.3  17.6  42.6
      0.220  38941   147   8280    96    82.4   99.8    1.7  17.5  39.5
      0.240  38928   185   8242   109    82.4   99.7    2.2  17.5  37.1
      0.260  38907   241   8186   130    82.5   99.7    2.9  17.4  35.0
      0.280  38895   298   8129   142    82.6   99.6    3.5  17.3  32.3
      0.300  38874   348   8079   163    82.6   99.6    4.1  17.2  31.9
      0.320  38842   431   7996   195    82.7   99.5    5.1  17.1  31.2
      0.340  38808   563   7864   229    82.9   99.4    6.7  16.8  28.9
      0.360  38746   747   7680   291    83.2   99.3    8.9  16.5  28.0
      0.380  38648   941   7486   389    83.4   99.0   11.2  16.2  29.2
      0.400  38534  1167   7260   503    83.6   98.7   13.8  15.9  30.1
      0.420  38361  1427   7000   676    83.8   98.3   16.9  15.4  32.1
      0.440  38152  1706   6721   885    84.0   97.7   20.2  15.0  34.2
      0.460  37911  2045   6382  1126    84.2   97.1   24.3  14.4  35.5
      0.480  37629  2393   6034  1408    84.3   96.4   28.4  13.8  37.0
      0.500  37309  2757   5670  1728    84.4   95.6   32.7  13.2  38.5
      0.520  36963  3166   5261  2074    84.5   94.7   37.6  12.5  39.6
      0.540  36585  3561   4866  2452    84.6   93.7   42.3  11.7  40.8
      0.560  36117  3925   4502  2920    84.4   92.5   46.6  11.1  42.7
      0.580  35622  4324   4103  3415    84.2   91.3   51.3  10.3  44.1
      0.600  35072  4709   3718  3965    83.8   89.8   55.9   9.6  45.7
      0.620  34481  5070   3357  4556    83.3   88.3   60.2   8.9  47.3
```

0.640	33850	5398	3029	5187	82.7	86.7	64.1	8.2	49.0
0.660	33211	5778	2649	5826	82.1	85.1	68.6	7.4	50.2
0.680	32611	6077	2350	6426	81.5	83.5	72.1	6.7	51.4
0.700	31932	6336	2091	7105	80.6	81.8	75.2	6.1	52.9
0.720	31277	6628	1799	7760	79.9	80.1	78.7	5.4	53.9
0.740	30669	6846	1581	8368	79.0	[illegible]	81.2	4.9	55.0
0.760	30070	7050	1377	8967	78.2	77.0	83.7	4.4	56.0
0.780	29415	7237	1190	9622	77.2	75.4	85.9	3.9	57.1
0.800	28720	7412	1015	10317	76.1	73.6	88.0	3.4	58.2
0.820	28021	7565	862	11016	75.0	71.8	89.8	3.0	59.3
0.840	27296	7746	681	11741	73.8	69.9	91.9	2.4	60.3
0.860	26505	7862	565	12532	72.4	67.9	93.3	2.1	61.4
0.880	25606	7965	462	13431	70.7	65.6	94.5	1.8	62.8
0.900	24624	8077	350	14413	68.9	67.1	95.8	1.4	64.1
0.920	23504	8153	274	15533	66.7	60.2	96.7	1.2	65.6
0.940	22172	8235	192	16865	64.1	56.8	97.7	0.9	67.2
0.960	20569	8304	123	18468	60.8	52.7	98.5	0.6	69.0
0.980	18085	8356	71	20952	55.7	46.3	99.2	0.4	71.5
1.000	149	8427	0	38888	18.1	0.4	100.0	0.0	82.2

```
;
DATA ONE; SET THRESH1F;
ESUM1F=FPOS+FNEG;
IF SEN EQ 0.0 THEN DELETE;
IF SPEC EQ 0.0 THEN DELETE;
PROC SORT; BY ESUM1F;
PROC PRINT; VAR LEVEL ESUM1F E1 E2 E3 E4;
```

3. System-Wide Model Program and Output

```
OPTIONS LS=80;
PROC LOGISTIC DATA=SASIN.FINAL;
MODEL YT/TOT=HOURS RCS ECC INCS ALT ALT2 LONG
LONG2 PER LAMBDA SOLF
LAMBDA2 LLMCROSS LACROSS ALMCROSS A3CROSS
GEOMDAY / CTABLE SELECTION=B SLSTAY=.2 FAST;
OUTPUT OUT=OUT1 P=PHAT;
DATA TWO; SET OUT1;
IF PHAT LE .38 THEN DELETE;
```

THE SAS SYSTEM
The LOGISTIC Procedure

Data Set: SASIN.FINAL
Response Variable (Events): YT
Response Variable (Trials): TOT
Number of Observations: 47464
Link Function: Logit

Response Profile

Ordered Value	Binary Outcome	Count
1	EVENT	51139
2	NO EVENT	160026

Backward Elimination Procedure

Step 0. The following variables were entered:

INTERCPT	HOURS	RCS	ECC	INCS	ALT	ALT2
LONG	LONG2	PER	LAMBDA	SOLF	LAMBDA2	LLMCROSS
LACROSS	ALMCROSS	A3CROSS	GEOMDAY			

Criteria for Assessing Model Fit

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
AIC	233793.20	190912.62	.

Criterion	Intercept Only	Intercept and Covariates	Chi-Square for Covariates
SC	233803.46	191097.31	.
-2 LOG L	233791.20	190876.62	42914.576 with 17 DF (p=0.0001)
Score	.	.	30587.645 with 17 DF (p=0.0001)

NOTE: No (additional) variables met the 0.2 significance level for removal from the model.

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Standardized Estimate	Odds Ratio
INTERCPT	1	-23.0699	1.1511	401.6573	0.0001	.	0.000
HOURS	1	-0.0553	0.000865	4088.3256	0.0001	-0.205307	0.946
RCS	1	0.0599	0.000975	3778.3616	0.0001	0.294000	1.062
ECC	1	-2.0058	0.0752	710.6047	0.0001	-0.235622	0.135
INCS	1	-0.5644	0.0205	757.7648	0.0001	-0.108204	0.569
ALT	1	0.00105	0.000034	947.5805	0.0001	3.821962	1.001
ALT2	1	-1.57E-8	3.92E-10	1612.4167	0.0001	-1.740138	1.000
LONG	1	-0.2402	0.0406	35.0673	0.0001	-3.630998	0.786
LONG2	1	0.00116	0.000198	34.2927	0.0001	3.436843	1.001
PER	1	0.00268	0.000114	554.8561	0.0001	0.314228	1.003
LAMBDA	1	1.4086	0.0738	364.7583	0.0001	3.549733	4.090
SOLF	1	0.00132	0.00021	39.6502	0.0001	0.021930	1.001
LAMBDA2	1	-0.00853	0.00105	66.1848	0.0001	-1.038614	0.992
LLMCROSS	1	-0.0005	0.000175	8.3502	0.0039	-0.246215	0.999
LACROSS	1	-8.06E-7	2.712E-7	8.8338	0.0030	-0.316710	1.000
ALMCROSS	1	-0.00003	1.105E-6	797.1808	0.0001	-2.814090	1.000
A3CROSS	1	6.272E-8	9.592E-9	42.7585	0.0001	0.608904	1.000
GEOMDAY	1	0.00510	0.00109	21.9561	0.0001	0.015994	1.005

The SAS System 2
14:56 Tuesday, September 15, 1992

The LOGISTIC Procedure

Association of Predicted Probabilities and Observed Responses

Concordant	= 79.4%	Somers' D	= 0.590
Discordant	= 20.4%	Gamma	= 0.592
Tied	= 0.2%	Tau-a	= 0.217
(8183569614 pairs)		c	= 0.795

Classification Table

Prob Level	Correct Event	Correct Non-Event	Incorrect Event	Incorrect Non-Event	Percentages Correct	Sensi-tivity	Speci-ficity	False POS	False NEG
0.000	51139	0	16E4	0	24.2	100.0	0.0	75.8	.
0.020	50834	26781	133E3	305	36.8	99.4	16.7	72.4	1.1
0.040	50531	37749	122E3	608	41.8	98.8	23.6	70.8	1.6
0.060	50048	46465	114E3	1091	45.7	97.9	29.0	69.4	2.3
0.080	49490	53435	107E3	1649	48.7	96.8	33.4	68.3	3.0
0.100	48820	59575	1E5	2319	51.3	95.5	37.2	67.3	3.7
0.120	48079	65701	94325	3060	53.9	94.0	41.1	66.2	4.5
0.140	47246	71394	88632	3893	56.2	92.4	44.6	65.2	5.2
0.160	46392	76828	83198	4747	58.4	90.7	48.0	64.2	5.8
0.180	45540	81916	78110	5599	60.4	89.1	51.2	63.2	6.4
0.200	44492	87572	72454	6647	62.5	87.0	54.7	62.0	7.1
0.220	43396	93092	66934	7743	64.6	84.9	58.2	60.7	7.7
0.240	42064	98720	61306	9075	66.7	82.3	61.7	59.3	8.4
0.260	40609	105E3	55401	10530	68.8	79.4	65.4	57.7	9.1
0.280	38862	11E4	49537	12277	70.7	76.0	69.0	56.0	10.0
0.300	36825	116E3	43894	14314	72.4	72.0	72.6	54.4	11.0
0.320	34727	122E3	38458	16412	74.0	67.9	76.0	52.5	11.9
0.340	32310	126E3	33605	18829	75.2	63.2	79.0	51.0	13.0
0.360	29484	131E3	28874	21655	76.1	57.7	82.0	49.5	14.2
0.380	26721	135E3	24853	24418	76.7	52.3	84.5	48.2	15.3
0.400	23754	139E3	20949	27385	77.1	46.4	86.9	46.9	16.5
0.420	20800	143E3	17250	30339	77.5	40.7	89.2	45.3	17.5
0.440	18087	146E3	14315	33052	77.6	35.4	91.1	44.2	18.5
0.460	15738	148E3	11529	35401	77.8	30.8	92.8	42.3	19.3
0.480	13631	151E3	9320	37508	77.8	26.7	94.2	40.6	19.9
0.500	11478	153E3	7140	39661	77.8	22.4	95.5	38.3	20.6
0.520	9524	155E3	5331	41615	77.8	18.6	96.7	35.9	21.2
0.540	7639	156E3	3867	43500	77.6	14.9	97.6	33.6	21.8

0.560	6065	157E3	2807	45074	77.3	11.9	98.2	31.6	22.3
0.580	4760	158E3	2002	46379	77.1	9.3	98.7	29.6	22.7
0.600	3621	159E3	1471	47518	76.8	7.1	99.1	28.9	23.1
0.620	2835	159E3	1209	48304	76.6	5.5	99.2	29.9	23.3
0.640	2336	159E3	1002	48803	76.4	4.6	99.4	30.0	23.5
0.660	1830	159E3	827	49309	76.3	3.6	99.5	31.1	23.6
0.680	1483	159E3	704	49656	76.2	2.9	99.6	32.2	23.8
0.700	1175	159E3	656	49964	76.0	2.3	99.6	35.8	23.9
0.720	920	159E3	604	50219	75.9	1.8	99.6	39.6	24.0
0.740	694	159E3	564	50445	75.8	1.4	99.6	44.8	24.0
0.760	561	16E4	516	50578	75.8	1.1	99.7	47.9	24.1
0.780	473	16E4	471	50666	75.8	0.9	99.7	49.9	24.1
0.800	418	16E4	448	50721	75.8	0.8	99.7	51.7	24.1
0.820	320	16E4	416	50819	75.7	0.6	99.7	56.5	24.2
0.840	296	16E4	392	50843	75.7	0.6	99.8	57.0	24.2
0.860	268	16E4	372	50871	75.7	0.5	99.8	58.1	24.2
0.880	248	16E4	337	50891	75.7	0.5	99.8	57.6	24.2
0.900	204	16E4	271	50935	75.8	0.4	99.8	57.1	24.2
0.920	178	16E4	249	50961	75.7	0.3	99.8	58.3	24.2
0.940	143	16E4	219	50996	75.7	0.3	99.9	60.5	24.2
0.960	103	16E4	173	51036	75.7	0.2	99.9	62.7	24.2
0.980	59	16E4	124	51080	75.8	0.1	99.9	67.8	24.2
1.000	0	16E4	0	51139	75.8	0.0	100.0	.	24.2

C. FINAL DATA SET USED IN ANALYSIS

One line of final data for 4 May 1992 set used in the analysis is provided.

DATE	TIME	SATCAT	RCS	ECC	INC	ALT	LONG	PER	Y1	Y2	Y3	Y4	Y5	Y6
920504	110128.767	5	.050	.1859	.826635	3710.84	144.60	133.2	1	1	1	1	0	0

LAMBDA	SOLF	ASOLF	GEOMD	GEOM
25.674	135	180	013	012

LIST OF REFERENCES

1. Schaaf, Steven F., *NAVSPASUR Sensor Performance Study,* Master's Thesis, Naval Postgraduate School, Monterey California, September, 1991.

2. *Random House Webster's College Dictionary*, p. 423, Random House, Inc., New York, 1990.

3. Naval Space Surveillance Command Internal Technical Report, *The Solar Flux and Geomagnetic Index*, by Stuart Boehmer, pp. 1-3, 3 July 1989.

4. Tascione, Thomas, F., *Introduction to the Space Environment*, p.45, Orbit Book Company, Malabar, Florida, 1988.

5. Allnutt, J.E., *Satellite-to-Ground Radio Wave Propagation, Theory, Practice and System Impact at Frequencies above 1 GHz,* pp. 59-65, Peter Peregrinus Ltd., London, United Kingdom, 1989.

6. Telephone conversation between Smith, Robin, (F80Q), Naval Space Surveillance Command and the author, 14 August 1991.

7. Wight, Randy, L., *SS3001 Military Applications of Space,* class notes presented to SS3001 class, Monterey, California, 23 September, 1991.

8. Weisberg, Sanford, *Applied Linear Regression, Second Edition,* pp. 267-270, John Wiley & Sons, Inc., New York, 1985.

9. *SAS/STAT User's Guide, Version 6, Fourth Edition Volume 2,* pp. 1071-1126, SAS Institute, Inc., Cary, North Carolina, 1990.

10. *SAS User's Guide, Basics, Version 5 Edition*, SAS Institute, Inc., Cary, North Carolina, 1990.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defence Technical Information Center Cameron Station Alexandria, Virginia 22314-6145	2
2.	Library, Code 52 Naval Postgraduate School Monterey, California 93943-5100	2
3.	Professor So Young Sohn OR/Sh Naval Postgraduate School Monterey, California 93943-5000	1
4.	Professor Dan C. Boger Code As/Bo Naval Postgraduate School Monterey, California 93943-5000	1
5.	Naval Space Surveillance Center 80 Department Dahlgren, Virginia 22448-5180	1